Vol. 32, No. 1 February, 1948

Journal of Applied Psychology

EDITED BY: DONALD G. PATERSON, UNIVERSITY OF MINNESOTA

Consulting Editors

PAUL S. ACHILLES, *Psychological Corporation;* WALTER V. BINGHAM, *A.G.O., War Department;* HAROLD E. BURTT, *Ohio State University;* ARTHUR I. GATES, *T. C. Columbia University;* JOHN G. JENKINS, *University of Maryland;* IRVING LORGE, *T. C. Columbia University;* QUINN MCNEMAR, *Stanford University;* WILLARD C. OLSON, *University of Michigan;* JAMES P. PORTER, *Swarthmore, Pennsylvania;* EDWARD K. STRONG, JR., *Stanford University;* MORRIS S. VITELES, *University of Pennsylvania;* JOSEPH ZUBIN, *N. Y. Psychiatric Institute.*

Table of Contents

Consistency of Voting by Our Congressmen: D. R. BRIMHALL AND A. S. OTIS..... 1
Job Evaluation of Nonacademic Work at the University of Illinois: A. M. JONES.. 15
Accident Proneness Among Street Car Motormen and Motor Coach Operators: C. W. BROWN AND E. E. GHISELLI 20
The Guilford-Zimmerman Aptitude Survey: J. P. GUILFORD AND W. S. ZIMMERMAN 24
Veterans' Scores on the Purdue Peyboard Test: J. K. STRANGE AND A. Q. SARTAIN 35
High School Norms for the Grove Modification of the Kent-Shakow Formboard Series: R. C. WYLIE, A. W. WILSON AND W. R. GROVE..... 41
Visual Skills and Labor Turnover: N. C. KEPHART..... 51
The Validity and Reliability of Heterophoria Scores Yielded by Three Commercial Optical Devices: J. H. SULZMAN, E. B. COOK, N. R. BARTLETT..... 56
Construction and Use of Weighted Check List Rating Scales for Two Industrial Situations: E. B. KNAUFT..... 63
Communication Between Management and Workers: D. G. PATERSON AND J. J. JENKINS 71
College Students' Opinions of Radio Advertising: G. P. LAYBOURN AND H. P. LONGSTAFF 81
Opinion Polling with Mark-Sensed Punch Cards: N. L. GAGE AND H. H. REMMERS 88
News Notes..... 92
Book Reviews..... 94
New Books, Monographs, and Pamphlets..... 102

Published Bi-monthly by The American Psychological Association, Inc.
Prince and Lemon Sts., Lancaster, Pa., and
1515 Massachusetts Ave., NW, Washington 5, D. C.

Entered as second-class matter, August 19, 1943, at the post office at Lancaster, Pa., under the act of March 3, 1879
Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. and R., authorized October 10, 1947

Copyright, 1948, by The American Psychological Association, Inc.

Journal of Applied Psychology

The *Journal of Applied Psychology* is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $6.00, single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Avenue, Northwest, Washington 5, D. C.

The character of the articles desired and the policy of this Journal with respect to subject matter is as follows: Reports of *original investigations* in any field of applied Psychology (except clinical and consulting psychology) are preferred. An occasional descriptive or theoretical article, however, will be accepted if it deals with some phase of applied psychology in a distinctive manner. The policy, however, is to favor papers dealing with quantitative investigations that would be of direct value to applied psychologists working in the following fields: Vocational diagnosis and occupational guidance, educational diagnosis, prediction and guidance at the secondary school level and higher, personnel selection, training, placement, transfer and promotion in business, industry and government service including the armed forces, training supervisors in business, industry and government service, problems of illumination, ventilation, fatigue, etc. in business and industry, job analysis, description, and evaluation, employee morale measurement, applied social and political psychology (such as opinion surveys conducted by The Psychological Corporation), market research and advertising. Articles under 500 words may be accepted. Maximum limit would be 16,000 words. The average would probably be in the neighborhood of 4,000 words. In order to reduce lag of publication, adherence to the rule of "brevity consistent with clarity" is encouraged.

With rare exceptions, articles are published in the order of their receipt.

The policy of "early publication," the author paying complete costs, is followed. The scheduled 80 pages per issue is thereby increased by the corresponding amount; thus the "early publication" of an article is a direct contribution to the subscribers without handicap to those authors whose articles are accepted in their regular turn. As a matter of fact, this policy directly benefits authors whose articles are published in their regular turn since it is an important means of cutting down lag in publication. It should be understood, however, that lag will necessarily vary because manuscripts are not received in an even flow throughout the year.

Articles submitted for publication should be sent to Donald G. Paterson, Editor, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota.

Authors should adhere to the conventions as described by J. E. Anderson and W. L. Valentine. The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. All manuscripts should be typed double-spaced throughout, including tables, footnotes, and references. A reprint will be loaned to any prospective author desiring same.

PUBLICATIONS OF

The American Psychological Association, Inc.

AMERICAN PSYCHOLOGIST

Editor: DAEL WOLFLE, *American Psychological Association*

Contains all official papers of the Association and articles concerning psychology as a profession; monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

Editor: GORDON W. ALLPORT, *Harvard University*

Contains original contributions in the field of abnormal and social psychology, reviews, and case reports; quarterly.
Subscription: $5.00 (Foreign $5.25). Single copies, $1.50.

JOURNAL OF APPLIED PSYCHOLOGY

Editor: DONALD G. PATERSON, *University of Minnesota*

Contains material covering applications of psychology to business, industry, education, etc.; bi-monthly.
Subscription: $6.00 (Foreign $6.50). Single copies, $1.25.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY

Editor: CALVIN P. STONE, *Stanford University*

Contains original contributions in the field of comparative and physiological psychology; bi-monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

JOURNAL OF CONSULTING PSYCHOLOGY

Editor: LAURANCE F. SHAFFER, *Teachers College, Columbia University*

Contains articles in the field of clinical and consulting psychology, counseling and guidance; bi-monthly.
Subscription: $5.00 (Foreign $5.50). Single copies, $1.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editor: FRANCIS W. IRWIN, *University of Pennsylvania*

Contains original contributions of an experimental character; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

Editor: C. M. LOUTTIT, *University of Illinois, Galesburg, Illinois*

Contains noncritical abstracts of the world's literature in psychology and related subjects; monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $.75.

PSYCHOLOGICAL BULLETIN

Editor: LYLE H. LANIER, *Vassar College*

Contains critical reviews of books and articles and critical and analytical summaries of psychological fields or subject matter; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED

Editor: HERBERT S. CONRAD, *College Entrance Examination Board, Princeton*

Contains longer researches and laboratory studies which appear as units; published at irregular intervals at a cost to author of about $2.50 a page; author receives 150 copies gratis.
Subscription: $6.00 per volume of about 350 pages (Foreign $6.30). Single copies, price varies according to size.

PSYCHOLOGICAL REVIEW

Editor: CARROLL S. PRATT, *Princeton University*

Contains original contributions of a theoretical nature; bi-monthly.
Subscription: $5.50 (Foreign $5.75). Single copies, $1.00.

Subscriptions are payable in advance and are terminated at expiration.
Make checks payable to the American Psychological Association, Inc.

Subscriptions, orders, and other business communications should be sent to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 MASSACHUSETTS AVENUE N.W., WASHINGTON 5, D.C.

The Basic Pictorial Material of Contemporary Psychology, Presented —for the first time—in Book Form

PSYCHOLOGICAL ATLAS

By DAVID KATZ

Professor of Psychology, Stockholm University

"PROFESSOR David Katz has given us a most valuable book. Physically, we can now cross the Atlantic in a number of hours. But, culturally, America and Europe tend to remain separate, if not to drift farther apart. Nowhere is this situation more obvious than it is in psychology. Under these circumstances, we must be grateful to Professor Katz for showing us by means of many impressive pictures how colorful and varied the interests of psychologists are in Europe."—*WOLFGANG KÖHLER, Research Professor of Philosophy and Psychology, Swarthmore College.*

The volume covers all major branches of psychology and allied subjects, carrying illustrative material in: NEUROLOGY, OPTICS, GESTALT-PSYCHOLOGY, CHARACTEROLOGY, TYPOLOGY, PSYCHOLOGICAL TESTING, PSYCHOLOGY OF CHILDREN AND ADOLESCENTS, PSYCHOLOGY OF THE HANDICAPPED, MEDICAL PSYCHOLOGY, PSYCHIATRY, APPLIED PSYCHOLOGY, EXPERIMENTAL PSYCHOLOGY, ANIMAL PSYCHOLOGY, REHABILITATION AND VOCATIONAL GUIDANCE.

$5.00

At Your Bookstore, or Use Handy Order Coupon:

LIMITED EDITION
ORDER NOW!

PHILOSOPHICAL LIBRARY, *Publishers*
15 East 40th St., Dept. 186, New York 16, N. Y.

Please send me_____copies of the PSYCHOLOGICAL ATLAS, at $5.00 per copy. Enclosed are $_____.

NAME.....

ADDRESS.....

Journal of Applied Psychology

Vol. 32, No. 1 February, 1948

Consistency of Voting by Our Congressmen

Dean R. Brimhall and Arthur S. Otis

Washington, D. C.

Can we predict from the past record of a congressman how he will vote in the future? The answer is yes—to a very marked degree.

This answer is based on a careful study of the votes cast by 512 congressmen on important propositions during the past five years.

Let us say that a congressman's position on a scale from *liberal* to *conservative* is represented by assigning him a scale value from 1 to 7, the value 1 representing the most liberal position and 7 the most conservative position, as shown in Figure 1. Then we may say, briefly, that if each

FIG. 1. Showing seven positions along a scale from liberal to conservative.

congressman is given a scale value each year to represent his position on the scale for that year, there are 46 chances in 100 that his scale value will not change from one year to the next; there are 83 chances in 100 that his scale value will not change more than 1 unit from one year to the next; and there are 95 chances in 100 that his scale value will not change more than 2 units. The maximum change possible would be 6 units.

Source of the Data

The data upon which the study was based were reported in the *New Republic*. In various past issues the vote of each congressman is given for each of a number of important propositions on which a roll call was had.

In the case of each proposition the vote of each congressman was recorded as follows:

\+ = a "progressive" vote, according to the opinion of the editors of the *New Republic*;
− = an "anti-progressive" vote;
0 = absence or "paired."

For example the votes of the Arkansas senators on the 10 propositions reported on August 4, 1947, were recorded as shown in Table 1.

Table 1

Manner of Reporting Votes

Proposition	1	2	3	4	5	6	7	8	9	10
Arkansas										
Fulbright	+	−	+	+	+	0	+	+	−	+
McClellan	−	−	+	−	+	−	+	+	−	+

The table shows that in the opinion of the editors, Senator Fulbright voted "progressively" on Proposition 1, voted "anti-progressively" on Proposition 2, voted "progressively" on Propositions 3, 4, and 5, was absent or paired on Proposition 6, etc.

For convenience the data were grouped and designated as shown in Table 2.

Table 2

Source of Data

Period	Dates of Reports	Number of Propositions		Designation Adopted
		Senate	House	
I	May 8, 1944	18	18	"1944"
II	Feb. 4, 1945	5	6	"1945"
	Feb. 11, 1946	10	10	
III	Sept. 23, 1946	14	15	"1946"
IV	Aug. 4, 1947	10	10	"1947"
	Total	57	59	

Since the propositions reported upon were not the same for the House and Senate, a separate study was made for each legislative body. The study of the voting of senators is described first.

"Progressiveness" Expressed as a Per Cent

First, a percentage rating was obtained for each senator for each period by finding the number of votes cast by a senator during that period and then finding what percentage of these were "progressive" votes.

For example, if, during Period I, a senator voted on 15 of the 18 propositions and voted "progressively" on 12 of these 15, his percentage ratings for that period was 12/15, or 80%.

Assigned scale values. The percentage rating of each senator for each of the four periods was then converted into a scale value according to Table 3.

Table 3

Scale Values

Highest seventh of the percentages	1
Next highest seventh percentages	2
Next highest seventh percentages	3
Next highest seventh percentages	4
Next highest seventh percentages	5
Next highest seventh percentages	6
Lowest seventh of the percentages	7

To assign these scale values it was necessary to find the distribution of percentage ratings and so arrange them in order.

For purposes of assigning these scale values, use was made of the percentage ratings of only those senators who were present in all four periods. There were 57 such senators, as may be determined from Table 4.

The highest seventh of the 57 percentage ratings in Period I were, respectively, 100%, 100%, 94%, 94%, 94%, 94%, 93%, and 92%. These were therefore assigned the scale value 1. The percentages constituting the next highest seventh were 85%, 81%, 76%, 76%, 76%, 75%, 71%, and 63%. These were assigned the scale value 2. Those constituting the next seventh were assigned the scale value 3, and so on.

The scale values for the four periods for the various senators are shown in Table 4. This table contains the names of only those senators who were present during at least two of the four periods. The starred senators are Republicans.

In order to present a single comprehensive scale value for each such senator, Table 4 shows also the *average* scale value for each senator who was present in Period IV.

Consistency in Voting

A glance at Table 4 will show that there is a marked degree of consistency in the scale values. Note for example how Hayden of Arizona got 2, 2, 2, 2, for the four periods.

To obtain a measure of consistency of voting, each scale value was compared with the one preceding, to see whether it was the same or whether it differed by one unit or two units or three units.

Thus, Hill of Alabama shifted first from 3 to 2 (one unit of change), then repeated the 2 (zero units of change) and then shifted from 2 to 1 (one unit). In this way the 218 shifts were noted. They are shown in Table 5.

Table 4
Scale Values of Senators

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Alabama					
Hill	3	2	2	1	2
Arizona					
Hayden	2	2	2	2	2
McFarland	2	2	3	3	3
Arkansas					
Fulbright			3	2	2
McClellan	6	4	4	3	4
California					
Downey	1	1	1	2	1
*Knowland			3	4	4
Colorado					
Johnson	3	3	2	3	3
*Millikin	7	7	7	5	6
Connecticut					
McMahon			1	2	2
Delaware					
*Buck	7	7	6	7	7
Tunnell	1	2	1		
Florida					
Andrews	3	5	5		
Pepper	1	1	1	1	1
Georgia					
George	5	4	5	4	4
Russell	4	2	4	3	3
Idaho					
Taylor			1	1	1
Illinois					
*Brooks	6	6	6	7	6
Lucas	2	2	2	2	2
Indiana					
*Capehart			5	7	6
*Willis	7	5	5		
Iowa					
*Hickenlooper			5	6	6
*Wilson	6	6	4	5	5
Kansas					
*Capper	4	4	6	6	6
*Reed	6	6	6	6	6
Kentucky					
Barkley	1	1	2	1	1
Louisiana					
Ellender	2	3	3	4	3
Overton	6	3	4	4	4
Maine					
*Brewster	4	6	5	7	6
*White	5	7	7	7	7
Maryland					
Radcliffe	2	4	4		
Tydings	5	4	5	3	4
Massachusetts					
*Saltonstall			3	5	4
Walsh	4	2	2		
Michigan					
*Ferguson	5	5	3	5	5
*Vandenberg	3	5	4	5	4
Minnesota					
*Ball	2	4	4	4	4
*Shipstead	4	4	4		
Mississippi					
Bilbo	4	4	5	7	5
Eastland	5	4	5	3	4
Missouri					
*Donnell			3	5	4
Montana					
Murray	1	1	1	1	1
Wheeler	5	4	3		
Nebraska					
*Butler	5	6	6	6	6
*Wherry	7	5	6	6	6
Nevada					
McCarran	3	4	3	4	4

Table 4—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
New Hampshire					
*Bridges	5	6	7	6	6
*Tobey	4	3	3	3	3
New Jersey					
*Hawkes	7	7	7	6	7
*Smith		4	4	5	4
New Mexico					
Chavez	4	2	1	2	2
Hatch	2	2	2	3	2
New York					
Mead	1	1	1		
Wagner	1	1	1	1	1
North Carolina					
Bailey	5	6	5		
Hoey			4	4	4
North Dakota					
*Langer	2	3	3	2	2
*Young			6	5	5
Ohio					
*Taft	5	5	4	7	5
Oklahoma					
*Moore	7	7	7	6	7
Thomas	3	4	4	2	3
Oregon					
*Cordon		5	4	7	6
*Morse			2	2	2
Pennsylvania					
Guffey	1	1	1		
Myers			1	2	2
Rhode Island					
Gerry	6	5	6		
Green	1	1	1	1	1
South Carolina					
Johnston			4	2	3
Maybank	3	3	4	3	3
South Dakota					
*Bushfield	7	7	7	6	7
*Gurney	4	6	7	5	5
Tennessee					
McKellar	5	4	3	5	4
Stewart	3	4	4	5	4
Texas					
Connally	5	5	6	4	5
O'Daniel	6	7	6	6	6
Utah					
Murdock	1	1	2		
Thomas	1	2	1	1	1
Vermont					
*Aiken	3	2	2	3	3
*Austin	3	5	4		
Virginia					
Byrd	6	6	4	5	5
Washington					
Magnuson		2	1	2	2
Mitchell		1	1		
West Virginia					
Kilgore	1	1	1	1	1
*Revercomb	7	6	4	6	6
Wisconsin					
LaFollette	2	2	1		
*Wiley	4	5	6	7	6
Wyoming					
O'Mahoney	2	3	3	2	2
*Robertson	7	7	7	6	7

Table 5 is read as follows. There were 90 instances in which the shift was zero units; that is, in which the scale value remained the same as it was the year before. These 90 cases were 41% of the 218 shifts. There were 91 instances (42% of the 218) in which the shift in scale value was only one unit. The 41% and 42% together made 83% of cases in which

Table 5

Shifts in Scale Values of Senators from One Period to the Next

Amount of Shift	No. of Cases	Per Cent	Sub-total
0 units	90	41	41
1 unit	91	42	83
2 units	34	16	99
3 units	3	1	100
Total	218	100	

the shift was not over one unit. There were 34 instances (16% of the 218) in which the shift was two units, making 99% of cases in which the shift was not over 2 units. In only 3 cases (about 1%) were the shifts as much as 3 units, as from 4 to 1.

The following are the three senators who shifted 3 units: Overton of Louisiana, 6 to 3; Taft of Ohio, 4 to 7; and Cordon of Oregon, 4 to 7.

The following senators maintained the same scale value throughout the four periods:

Hayden, Arizona,	2 2 2 2
Pepper, Florida,	1 1 1 1
Lucas, Illinois,	2 2 2 2
Reed, Kansas,	6 6 6 6
Murray, Montana,	1 1 1 1
Wagner, New York,	1 1 1 1
Green, Rhode Island,	1 1 1 1
Kilgore, W. Va.,	1 1 1 1

The Representatives

A comparable study was made of the consistency of voting of the representatives.

The scale values of those representatives who were present during at least two periods are shown in Table 8, near the end of the article, together with an average scale value for those who were present in the last period. The starred representatives are Republicans.

The degree of consistency of voting by representatives is shown in Table 6 which is comparable to Table 5.

Table 6 is read as follows: Of the 1076 shifts made by the representatives (in most cases 3 shifts by each), 504, or 47%, were zero units—no shift at all; 391, or 36%, were shifts of one unit, making 83% of cases in which the shift was not over one unit; 127, or 12%, were shifts of two units, etc.

Table 6

Shifts of Scale Values of Representatives from One Period to the Next

Amount of Shift	No. of Cases	Per Cent	Sub-total
0 units	504	47	47
1 unit	391	36	83
2 units	127	12	95
3 units	46	4.2	99.2
4 units	7	0.7	99.9
5 units	1	0.1	100.0
Total	1076	100.0	

The representative who made a shift of 5 units was Winstead of Mississippi, who shifted from 7 in "1946" to 2 in "1947."

Table 7 shows the data of Tables 5 and 6 combined; that is, for congressmen of both houses together.

Table 7

Shifts in Scale Values of Senators and Representatives Together

Amount of Shift	No. of Cases	Per Cent	Sub-total
0 units	594	46	46
1 unit	482	37	83
2 units	161	12	95
3 units	49	4	99
4 units	7	0.5	
5 units	1	0.1	

Table 7 shows that, as stated at the outset, there are 46 chances in 100 that a congressman's scale value will not change at all from any given year to the next; there are 83 chances in 100 that his scale value will not change more than one unit; and there are 95 chances in 100 that his scale value will not change more than 2 units.

The chances are less than 1 in 100 that a congressman's scale value will change more than three units out of a possible 6 units of change; and there is less than 1 chance in 1000 that a congressman's scale value will change 5 units.

Reversing the Scale Values

It goes without saying, no doubt, that the consistency of congressmen is the same whether we rate the most liberal congressman 1 and the most conservative 7, or whether we rate the most conservative 1 and the most liberal 7. A congressman whose scale value changed from 1 to 2 by the

Table 8

Scale Values of Representatives

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Alabama					
Andrews		5	5	2	4
Boykin	4	6	7	6	6
Grant	4	5	6	3	4
Hobbs	4	4	3	3	3
Jarman	3	3	3	3	3
Manasco	4	5	5	2	4
Patrick		1	1		
Rains		2	2	2	2
Sparkman	3	2	2		
Arizona					
Harless	1	2	2	2	2
Murdock	1	1	1	2	1
Arkansas					
Cravens	6	5	7	4	4
Gathings	5	7	6	6	6
Harris	3	2	3	2	2
Hays	2	2	2	2	2
Mills	3	3	4	3	3
Norrell	6	5	4	5	5
Trimble		3	2	2	2
California					
*Anderson	5	5	5	7	6
Douglas		1	1	1	1
Doyle		1	1		
Elliot	7	5	5	4	5
Engle	2	3	2	4	3
*Gearhart	7	7	5	7	6
Havenner		1	1	1	1
Healy		1	1		
*Hinshaw	6	4	3	6	5
Holifield	1	1	2	1	1
Izac	1	1	1		
*Johnson, J. L.	4	4	3	4	4
King	1	1	1	1	1
Lea	3	3	3	4	3
*McDonough		3	2	6	5
Miller		1	1	1	1
Outland	1	1	1		
Patterson		1	1		
*Phillips	5	7	6	6	6
Sheppard	3	2	2	2	2
Tolan	2	2	1		
Voorhis	1	1	2		
*Welch	2	2	2	2	2
Colorado					
*Chenoweth	7	6	6	4	6
*Gillespie		5	4		
*Hill	6	7	5	5	6
*Rockwell	6	7	6	4	6
Connecticut					
Geelan		1	1		
Kopplemann		1	1		
*Luce	3	3	2		
Ryter		1	1		
*Talbot	6	6	5		
Woodhouse		1	1		
Delaware					
Traynor		2	2		
Florida					
Cannon	3	6	3		
Hendricks	3	4	5	3	4
Peterson	3	3	4	3	3
Price	3	4	5	3	4
Rogers		4	5	3	4
Sikes	3	5	6	2	4
Georgia					
Brown	4	3	3	3	3
Camp	4	3	3	3	3
Cox	4	5	6	3	4
Gibson	5	6	6		
Pace	3	3	3	3	3
Peterson	4	4	4		
Tarver	4	3	5		
Vinson	4	2	3	3	3
Wood		4	7	4	5
Idaho					
*Dworshak	7	7	6		
White	2	3	3		

Table 8—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Illinois					
*Allen	6	7	6	6	6
*Arends	6	7	7	7	7
*Bishop	6	5	4	4	5
*Chipfield	6	7	7	6	6
*Church	6	7	7	4	6
Dawson	1	1	1	1	1
*Dirksen	6	4	4	5	5
Douglas		1	1		
Gordon	1	1	1	1	1
Gorski	1	1	1	1	1
*Howell	6	5	4	6	5
*Johnson	7	7	7	6	7
Kelly		2	2		
Link		1	1		
*McMillen		6	4	7	6
*Mason	7	7	7	6	7
O'Brien	2	1	1	2	2
Price		1	1	1	1
*Reed	7	7	6	7	7
Ressa		1	1		
Rowan	1	1	1		
Sabath	1	1	1	1	1
*Simpson	6	6	5	4	5
*Sumner	7	7	7		
*Vursell	6	7	6	6	6
Indiana					
*Gillie	6	7	6	7	7
*Grant	6	7	6	6	6
*Halleck	7	7	7	7	7
*Harness	7	7	7	6	7
*Johnson, N.	5	6	7	6	6
*LaFollette	4	2	1		
*Landis	5	7	6	6	6
Ludlow	2	2	1		
Madden	1	1	1	1	1
*Springer	6	6	6	7	6
*Wilson	6	5	5	4	5
Iowa					
*Cunningham	4	6	6	6	6
*Dolliver		6	5	4	5
*Gwynne	6	7	7	6	6
*Hoeven	4	7	6	6	6
*Jensen	7	7	7	5	6
*LeCompte	5	6	5	6	6
*Martin	5	7	6	4	6
*Talle	5	7	6	7	6
Kansas					
*Carlson	6	7	6		
*Cole		5	4	6	5
*Hope	6	7	4	5	5
*Rees	6	6	5	7	6
*Scrivner	7	6	6	6	6
*Winter	7	7	7		
Kentucky					
Bates	2	1	2	1	1
Chapman	3	3	3	3	3
Chelf		2	2	3	2
Clements		2	2	1	2
Gregory	3	3	2	2	2
May	7	4	3		
O'Neal	2	2	1		
*Robsion	6	6	5	7	6
Spence	1	1	1	1	1
Louisiana					
Allen	3	5	7	4	5
Brooks	3	3	5	4	4
Domengeaux	4	5	7	5	5
Hebert	2	6	7	4	5
Larcade	2	5	5	4	4
McKenzie	5	7	5		
Maloney	2	3	3		
Morrison	3	3	2	4	3
Maine					
*Fellows	7	7	6	5	6
*Hale	6	5	5	4	5
*Smith	5	2	2	4	3
Maryland					
Baldwin	5	3	3		
*Beall	5	5	5	7	6
D'Alesandro	1	2	2	2	2
Fallon		3	2	3	3
Roe		4	7		
Sasscer	2	3	2	2	2

Table 8—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Massachusetts					
*Bates	5	5	4	6	5
*Clason	7	5	3	4	5
Curley	2	2	1		
*Gifford	7	6	6	7	7
*Goodwin	7	7	6	4	6
*Herter	4	4	3	4	4
*Heselton		4	3	4	4
*Holmes	5	5	5		
Lane	1	1	2	2	2
McCormack	1	1	1	1	1
*Martin	6	6	6		
Philbin	2	4	3	2	3
*Rogers	7	7	4	6	6
*Wigglesworth	6	5	4	6	5
Michigan					
*Blackney	5	5	6	7	6
*Bradley	7	7	6	3	6
*Crawford	7	6	6	6	6
Dingell	1	1	1	1	1
*Dondero	6	6	6	6	6
*Engel	5	3	4	6	5
*Hoffman	7	7	7	5	6
Hook		1	1		
Jonkman	7	7	6	7	7
Lesinski	1	1	1	1	1
*Michener	6	6	4	7	6
O'Brien	2	1	1		
Rabaut	2	1	1		
Sadowski	1	1	1	1	1
*Shafer	7	6	6	5	6
*Wolcott	7	6	5	7	6
*Woodruff	7	7	7	7	7
Minnesota					
*Andersen	4	7	6	4	5
*Andresen	6	7	7	5	6
Gallagher		1	1		
*Hagen	3	5	5	5	5
*Judd	4	4	3	4	4
*Knutson	7	7	7	6	7
*O'Hara	7	7	7	6	7
*Pittenger	4	6	6		
Starkey		1	1		

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Mississippi					
Abernethy	3	5	6	3	4
Colmer	5	4	4	3	4
McGehee	5	6	7		
Rankin	3	7	7	3	5
Whitten	3	7	7	4	5
Whittington	3	3	4	2	3
Winstead	3	6	7	2	4
Missouri					
*Arnold	6	6	7	7	7
Bell	5	4	4	4	4
*Bennett	6	7	6	4	6
Cannon	3	1	1	2	2
Carnahan		1	1		
Cochran	1	1	1		
*Cole	5	6	5	6	6
*Ploeser	7	6	6	7	7
*Schwabe	5	7	7	4	6
*Short	7	7	7	5	6
Slaughter	4	4	5		
Sullivan		1	1		
Zimmerman	2	3	2	3	3
Montana					
*D'Ewart		6	5	5	5
Mansfield	1	2	2	2	2
Nebraska					
*Buffett	7	7	7	6	7
*Curtis	6	7	7	4	6
*Miller	5	7	7	5	6
*Stefan	4	7	7	6	6
Nevada					
Bunker		2	2		
New Hampshire					
*Adams		5	5		
*Merrow	6	4	5	4	5
New Jersey					
*Auchincloss	6	7	3	6	5
*Canfield	3	3	2	4	3
*Case		4	3	4	4
*Eaton	6	6	6	5	6

Table 8—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
*Hand		4	3	5	4
Hart	1	1	1	2	1
*Hartley	4	4	4	7	5
*Kean	5	4	3	4	4
Norton	1	1	1	1	1
*Sundstrom	6	5	4	6	5
*Thomas	7	7	6	5	6
*Towe	4	4	4	6	5
*Wolverton	3	3	3	5	4
New Mexico					
*Fernandez	2	3	2	3	3
New York					
*Andrews	5	6	5	4	5
*Baldwin	2	2	2		
Barry	2	2	2		
*Bennet		3	3		
Bloom	1	2	1	1	1
*Buck		7	5	4	5
Buckley	1	2	2	1	1
*Butler	5	5	3	4	4
Byrne	1	1	1	1	1
Celler	2	1	1	1	1
*Cole	5	7	7	6	6
Delaney, Jas.		1	1		
Delaney, John	1	1	1	2	1
*Elsaesser		6	4	5	5
*Fuller	7	7	6	7	7
*Gamble	7	6	4	5	5
*Gwinn		6	5	6	6
*Hall, E. A.	6	5	4	6	5
*Hall, L. W.	6	6	4	6	6
*Hancock	6	6	6		
Heffernan	1	1	1	2	1
*Kearney	3	5	4	6	5
Keogh	1	1	1	2	1
*Kilburn	7	7	5	6	6
*Latham		5	5	6	5
*LeFevre	5	6	6	7	6
Lynch	1	1	1	1	1
Marcantonio	1	1	1	1	1
O'Toole	1	1	1	1	1
Pfeifer	2	1	2	1	1
Powell		1	2	1	1
Quinn		1	1		
Rabin		1	1	1	1
Rayfiel		1	1	1	1
*Reed	7	7	7	7	7
Roe		2	2		
Rogers		1	1		
Rooney		1	1	2	1
*Sharp		6	4		
Somers	1	1	2	2	2
*Taber	7	7	7	7	7
*Taylor	5	4	4	5	5
Torrens		1	1		
*Wadsworth	7	7	7	5	6
North Carolina					
Barden	3	5	5	3	4
Bonner	2	3	3	3	3
Bulwinkle	2	2	3	4	3
Clark	3	2	2	3	3
Cooley	2	3	3	2	2
Doughton	3	3	4	3	3
Durham	4	3	3	3	3
Ervin		3	3		
Folger	1	2	2	1	1
Kerr	3	3	4	4	4
Weaver	2	2	3		
North Dakota					
*Lemke	3	3	4	3	3
*Robertson		6	6	6	6
Ohio					
*Bender	4	3	3	6	4
*Bolton	4	4	4	5	4
*Brehm	7	6	5	7	6
*Brown	7	7	7	5	6
*Clevenger	7	7	7	6	7
Crosser	1	1	1	1	1
*Elston	6	7	7	7	7
Feighan	1	1	1	1	1
Gardner		2	1		
*Griffiths	7	6	5	6	6
*Hess	5	7	7	7	7
Huber		1	1	1	1
*Jenkins	7	6	6	7	7

Table 8—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
*Jones	7	7	7	5	6
Kirwan	1	1	1	1	1
*Lewis	5	3	4	4	4
*McCowen	6	6	6	6	6
*McGregor	6	6	5	6	6
*Ramey	4	3	3	4	4
*Smith	7	7	7	4	6
Thom		2	1		
*Vorys	6	7	5	6	6
*Weichel	7	7	7	6	7
Oklahoma					
Boren	6	3	5		
Johnson	2	4	3	2	3
Monroney	2	2	2	2	2
*Rizley	7	7	7	6	7
*Schwabe		7	7	6	7
Stewart	4	5	5		
Stigler		3	2	2	2
Wickersham	3	3	3		
Oregon					
*Angell	3	4	3	2	3
*Ellsworth	5	6	5	6	6
*Stockman	6	5	6	6	6
Pennsylvania					
Barret		1	1		
Bradley	1	1	1		
*Brumbaugh	7	6	5		
*Campbell		5	5		
*Corbett		4	4	4	4
Eberharter	1	1	1	1	1
*Fenton	5	5	5	6	5
Flood		1	1		
*Fulton		2	2	4	3
*Gavin	6	6	5	7	6
*Gerlach	6	4	4	7	5
*Gillette	7	7	6	7	7
*Graham	6	5	5	7	6
Granahan		1	1		
Green		1	1		
*Gross	7	6	5	7	6
Hoch	1	1	1		
Kelley	1	1	1	1	1
*Kinzer	7	6	6		
*Kunkel	5	3	3	5	4
*McConnell	7	5	4	7	6
McGlinchey		1	1		
Morgan		1	1	1	1
Murphy		1	1		
*Rich	7	7	7	6	7
*Rodgers	6	6	6		
Sheridan	2	2	1		
*Simpson	7	7	6	7	7
*Tibbott	6	6	5	7	6
Walter	2	2	2	1	2
*Wolfenden	5	5	4		
Rhode Island					
Fogarty	1	2	2	1	1
Forand	1	1	1	1	1
South Carolina					
Bryson	3	3	3	3	3
Hare	3	4	4		
McMillan	3	4	4	2	3
Richards	2	3	3	3	3
Riley		3	3	3	3
Rivers	3	6	4	3	4
South Dakota					
*Case	4	6	4	4	4
*Mundt	5	7	7	6	6
Tennessee					
Cooper	2	3	3	3	3
Courtney	2	3	2	3	3
Davis	2	2	2	3	2
Earthman		4	6		
Gore	2	3	2	2	2
*Jennings	7	6	6	7	7
Kefauver	1	1	1	1	1
Murray	5	6	6	4	5
Priest	1	2	2	3	2
*Reece	7	7	5		
Texas					
Beckworth	3	3	3	3	3
Combs		2	2	1	2
Fisher	5	4	5	4	4

Table 8—*Continued*

	I	II	III	IV	
	1944	1945	1946	1947	Avg.
Gossett	4	5	5	4	4
Johnson, L. A.	3	3	3		
Johnson, L. B.	2	2	2	2	2
Kilday	6	3	6	4	5
Lanham	5	5	5		
Lyle		2	3	2	2
Mahon	3	3	5	3	3
Mansfield	3	3	4	1	3
Patman	2	2	2	3	2
Pickett		3	4	2	3
Poage	5	3	4	2	3
Russell	3	4	5		
Sumners	4	3	4		
Thomas	3	2	2	3	3
Thomason	3	3	2	2	2
West	7	7	7	4	6
Worley	3	5	4	3	4
Utah					
Granger	1	1	2	1	1
Robinson	1	1	1		
Vermont					
*Plumley	7	7	5	5	6
Virginia					
Bland	3	3	3	2	3
Burch	3	2	4		
Daughton		2	4		
Drewry	3	3	3	3	3
Flannagan	1	2	2	2	2
Gary		2	3	3	3
Robertson	4	4	4		
Smith	4	4	3	2	3
Washington					
Coffee	1	1	2		
Delacy		1	1		
*Holmes	6	4	3	6	5
*Horan	6	7	5	6	6
Jackson	1	1	1	1	1
Savage		1	1		
West Virginia					
Bailey		1	2		
*Ellis	7	7	7	5	6
Hedrick		1	2	2	2
Kee	1	1	1	1	1
Neely		1	1		
Randolph	3	2	2		
Wisconsin					
Biemiller		1	1		
*Byrnes		6	6	6	6
*Henry		5	4		
Hull	2	2	2	3	2
*Keefe	5	6	5	7	6
*Murray	4	5	4	7	5
*O'Konski	3	4	3	4	4
*Smith	6	5	4	6	5
*Stevenson	3	4	4	7	5
Wasielewski	3	2	2		
Wyoming					
Barrett	6	7	5	6	6

first method would merely change from 7 to 6 by the second method. The degree of consistency is the same in both cases.

Further Research Suggested

One of the main reasons for making and reporting this study is to stimulate further research along this line. Psychologists have investigated the consistency of behavior of clerks, salesmen, aircraft pilots, mechanics, taxi-cab drivers, and a host of other occupations. Yet so far as we know this is the first time a study of consistency has been made of the most important employees of democracy.

Further interesting researches would be to compare Republicans and Democrats either in their degrees of liberalism-conservatism or in their consistency of voting; to compare Southern Democrats with Northern Democrats; to investigate the consistency of voting in State legislatures, city councils, etc., and to compare scale values in liberalism-conservatism and other dimensions obtained in Congress with those previously had in State legislatures, etc.

Received October 8, 1947.
Early publication.

Job Evaluation of Nonacademic Work at the University of Illinois *

Alice May Jones

University of Illinois

Job evaluation is a systematic procedure which attempts to appraise the worth of a job, relatively considered with respect to other jobs in an organization, in direct proportion to its respective monetary value. J. E. Walters says, "It is the determination of the value of the job and not the worth of the employee doing that job" (6).

The history of job evaluation is replete with literature and requires no elaboration. The various aspects of this field, particularly the point system, are reviewed by J. E. Walters (6). Time-study methods are presented by Brash (3). The number of factors involved in a plan for evaluation of jobs is treated by Bradbury (2) and Viteles (5), and the relative weight each factor should bear is considered by Moore (4).

The review of literature on this subject indicates it is common practice for supervisors and specialists to evaluate various jobs in an organization, while it reveals employee-participation in the job evaluation plans of only two organizations; namely, the General Foods Corporation (7), and a midwestern manufacturing plant of paper specialties reported by Benge (1).

The purpose of this study is to describe the results of a plan of job evaluation at the University of Illinois which considered employee-participation, as well as the participation of supervisors and the personnel department.

Procedure

In the job evaluation plan which the University of Illinois adopted, nonacademic jobs requiring Civil Service appointment were evaluated on a point scale. By way of explanation, the Civil Service system in effect at the University of Illinois, which is entirely separate from the Civil Service system of the State of Illinois, contains 300 classifications,

* Submitted as Master's Thesis, under the direction of Professor T. W. Harrell and with the cooperation of the Non-academic Personnel Bureau, in partial fulfillment of the requirements for the degree of Master of Arts in Psychology in the Graduate School of the University of Illinois, 1946.

involving 2300 employees, 1500 of whom are included in the 93 classifications involved in this study. At the time of this study, ratings on the other 207 classifications, involving 800 employees, were not available. Since that time, these ratings have been submitted to the Office of Nonacademic Personnel, and that office is now in the process of studying them.

Regarding the classifications involved in this study, the Office of Nonacademic Personnel sent a copy of "Job Evaluation Work Sheet" to every supervisor of and employee on all jobs under Civil Service appointment. Each supervisor and employee was asked to assign numerical values to any given job, according to the points allocated for each of four major factors, separated into 14 minor factors, listed on the "Description and Value of Factors used in Job Evaluation."

The four major factors, each of which had a maximum point value, were "preparation for the job," 30 points; "personal qualifications," 30 points; "working conditions," 10 points; and "responsibilities required on the job," 60 points. Thus, 130 was the total number of points it was possible to assign to any one job classification.

Forty-two supervisors submitted ratings on 62 jobs, ranging in numerical value from 17 to 119 points, and employees of ten departments and committees submitted ratings on 64 jobs ranging in value from 13 to 103 points. Of the 93 classifications involved, supervisors and employees each submitted ratings for 33 of the same classifications, the other 60 jobs being appraised by either supervisors or employees. In addition, each of the 93 classifications was evaluated by the writer under the direction of Dr. Arlyn Marks of the Office of Nonacademic Personnel, which hereinafter are termed "personnel department" ratings. The classifications evaluated by the personnel department ranged from 16 to 90 points.

Three classifications were rated by ten or more supervisors; namely, "clerk-stenographer, junior" rated by 32 supervisors, "clerk-stenographer, senior" by 21 supervisors, and "clerk-typist, junior" by 10 supervisors. Ten classifications were each rated by two employees, the remaining 54 classifications bearing single employee ratings. In this connection, it should be mentioned that 41 of the 64 classifications were evaluated by a committee which represented the combined opinions of many employees.

Results and Discussion

The reliability of the point ratings submitted by 21 supervisors on the classification "clerk-stenographer, senior," was determined in the following manner: The ratings submitted by the 21 supervisors were listed at random, and the number of points assigned by each supervisor to each of the 14 factors involved were tabulated. The averages for each of the

14 factors for 10 supervisors were correlated with the averages for each of the 14 factors for the other 11 supervisors. The result was a correlation of .89 ± .04, and when corrected for small number of cases, the correlation became .88. (The correlations stated in this study are corrected correlations, provided the number of cases involved is less than 50. If the number of cases is over 50, the correlation has not been corrected.)

In the same manner as that for "clerk-stenographer, senior," the reliability of 32 supervisors' ratings (two groups of 16 supervisors' ratings) for "clerk-stenographer, junior," was computed to be .95 ± .02.

In order to ascertain the reliability of supervisors' point ratings on 30 different job classifications evaluated by two or more supervisors, total ratings for each job classification submitted by the various supervisors were listed randomly. The first two total ratings on each job were used in determining the reliability of such ratings, the first total rating on each job being an "X" and the second total rating on the same job being a "Y." The X's were correlated with the Y's, and a correlation coefficient of .50 ± .09 was computed.

The reliability of the employees' point ratings on 10 jobs, which were appraised by two or more employees, was determined in the same manner as that for the supervisors. A reliability coefficient of .59 ± .12 was found.

In order to ascertain the degree of agreement among employees, supervisors and personnel department, intercorrelations were found for ratings submitted by these three groups on the classifications involved. The ratings submitted by supervisors were averaged into a single rating for each job classification, and, likewise, the ratings submitted by employees were averaged into a single rating for each job classification. The ratings of the personnel department represented single ratings for each job.

On the 62 classifications rated by both supervisors and personnel department, the correlation was .86 ± .02. On the 64 classifications rated by both the employees and personnel department, the correlation was .85 ± .02. On the 33 classifications rated by both supervisors and employees, the correlation of their ratings was computed to be .89 ± .02. Considering the ratings of the personnel department on these 33 Civil Service classifications, the correlation between employees' and personnel department's ratings was .90 ± .02, and the correlation between the ratings of supervisors and those of the personnel department was .86 ± .03.

Algebraic differences between the point value ratings submitted by employees, supervisors and personnel department were determined to provide additional information regarding the relation between the evalu-

ations of these three groups. Considering the personnel department's ratings as a constant, the algebraic differences between the average rating of the employees and that of the personnel department on each of the 64 classifications rated by both totalled 181.5 points, and the algebraic differences between the average rating of supervisors and that of the personnel department on each of the 62 classifications rated by both totalled 264.3 points. This means that the ratings of supervisors, considered as a total, were higher than those of employees, and that there was closer agreement between the ratings of employees and personnel department on these classifications, than there was between supervisors and personnel department on these classifications.

The fact that the evaluations submitted by supervisors were higher than those submitted by employees for the same jobs may have been for one or more of three reasons: first, employees may rate their jobs conservatively because of modesty; second, if the supervisors assign high ratings to employees' jobs, this fact promotes their prestige and importance in their own jobs; and, third, supervisors may be anxious to seek the favor of their employees and so will assign high ratings to the employees' jobs.

As stated previously, it is a common practice for supervisors to rate jobs of their employees, but employees, themselves, are seldom asked to rate their own jobs, particularly when their ratings will be considered in a plan of job evaluation. This feeling of participation may have caused employees, more than supervisors, to be more conservative in their judgments of their jobs.

Summary and Conclusions

1. The results of this study indicate a high degree of agreement among the point ratings of employees, supervisors and personnel department on the classifications involved, in so far as the relative levels of jobs are concerned.

2. When algebraic differences were taken for the actual point ratings on the classifications involved, the evaluations of the employees agreed more closely with those of the personnel department, than did those of the supervisors.

3. So far as the writer has been able to ascertain, there has been no correlation published for the reliability between the ratings of employees and those of supervisors or management in similar studies. The correlations between the ratings of employees and supervisors, between employees and personnel department, and between supervisors and personnel department, are given in this study.

Received April 25, 1947.

References

1. Benge, E. J. Job evaluation in a paper plant. *Person. J.*, 1940, 19, 42–48.
2. Bradbury, K. F. Job evaluation analyzed. *Advanced Management*, 1940, 5, 16–20.
3. Brash, J. A. Time-study methods applied to job evaluation. *J. consult. Psychol.*, 1945, 9, 152–160.
4. Moore, H. Problems and methods of job evaluation. *J. consult. Psychol.*, 1944, 8, 90–99.
5. Viteles, M. S. A psychologist looks at job evaluation. *Personnel*, 1941, 17, 165–176.
6. Walters, J. E. *Personnel relations.* New York: The Ronald Press Co., 1945.
7. *Job evaluation:* Formal plans for determining basic pay differentials, Studies in Personnel Policy No. 25, National Industrial Conference Board, 1940.

Accident Proneness Among Street Car Motormen and Motor Coach Operators

Clarence W. Brown and Edwin E. Ghiselli

University of California

The term, accident proneness, has been used in two somewhat different meanings. In some instances the term is used to refer to the individual who is involved in many accidents of the same sort. Thus the operator of a public conveyance who has had many more accidents than the average of his fellow operators is termed accident prone. In this sense the term implies sheer quantity of accidents. In other cases accident proneness is used to describe the individual who has had accidents of several different kinds. Thus the individual who has had accidents at home, at work, and on the public highways is called accident prone.

While these two uses of the term, accident prone, are not wholly different, the second interpretation has certain implications not contained in the former. The facts of individual differences in accident rate is sufficient proof, if proof is needed, to substantiate the first viewpoint. Some persons do incur more accidents than do others. The second use of the term, however, requires statistical evidence in the form of correlations between rates of accidents of different kinds. If, for example, it were found that persons who had most accidents at work also were those who had most accidents at home, or if persons having most minor accidents also had most major accidents, then positive evidence would be at hand for supporting the notion of accident proneness as a general trait. Low or zero coefficients of correlation among different types of accidents would argue against such a concept.

While the evidence concerning the validity of the concept of accident proneness as a retention of liability under all circumstances is far from extensive, there is rather general agreement among investigators to the effect that the relationships among accidents of different kinds are low. Working with groups of factory workers, Newbold (1) found coefficients of correlation of the order of .20 to .30 between various types of accidents which occurred on the job and accidents which occurred at home. For various groups of laborers and craftsmen, Farmer and Chambers (2) obtained coefficients of correlation between major and minor accidents ranging from zero to about .35. A coefficient of the order of .20 to .25

might be considered representative of their data. For motor coach operators, Brown and Ghiselli (3) found a coefficient of correlation of .25 between collision and non-collision accidents.

One of the major limitations of the data collected to date on the generality of accident proneness is the scarcity of information on the reliability of the measures employed in measuring accident rate in the situations being compared. Numerous findings on the reliability of accident rates have been published when only individual accidents were under investigation. These reliability coefficients, of course, vary with the nature of the job involved and with the length of time considered. While they frequently approach zero they seldom rise above .80, most of them falling in the neighborhood of .30 to .60. It is apparent, then, that the coefficients of correlation between different types of accidents as reported by various investigators are considerably lower than the reliability coefficients of measures of accident rates. However, since data relative to the intercorrelations between different types of accidents and the reliabilities of those same types of accidents are not usually available, only very tentative conclusions can be drawn concerning the concept of accident proneness as a general trait. In an attempt to secure further evidence on this matter, the present investigation was undertaken.

Basic Data

The men considered herein are 59 trolley car motormen and 34 motor coach operators employed in a city transit system. All of these men had had two years experience on their jobs. The accident records taken for them covered a period of 18 months, and consisted of the entire year 1945 and the first half of 1946. All irregular cases, such as those who were under suspension or were absent without leave, etc., were eliminated.

The accidents of motormen were classified as follows: collisions with pedestrians, collisions with other trolley cars, collisions with motor vehicles, boarding and alighting accidents, and accidents aboard the car such as slips, stumbles, falls, etc., of passengers. These accidents were also grouped into the two major categories of collision and non-collision accidents. Since for the motor coach operators the rates for collisions with other than motor vehicles and for on-board accidents were very low, only the more inclusive grouping of collision and non-collision accidents was used.

Results

For all types of accidents, the reliability coefficients were estimated from the coefficient of correlation between the number of accidents on the odd and even months corrected by the Sperman-Brown formula. These

coefficients, given specifically in Tables 1 and 2, varied from .19 to .80, most of them being about .40.

In Table 1 are given the reliability coefficients, intercorrelations,and distribution constants for the various types of accidents of the streetcar

Table 1

Reliability Coefficients, Intercorrelations and Distribution Constants of Various Types of Accidents of 59 Trolley Car Motormen

	1	2	3	4	5	Mean	Standard Deviation
1. Collisions with pedestrians	*.46*	.10	−.11	.12	.02	.75	1.54
2. Collisions with trolley cars		*.19*	.22	.04	.17	.44	.65
3. Collisions with motor vehicles			*.42*	.03	.07	12.22	5.87
4. Boarding and alighting accidents				*.15*	.19	.42	.69
5. On-board accidents					*.65*	.27	.45

motormen. It will be noted that the coefficients of correlation between the various types of accidents are generally quite low, and one, indeed, is negative. Since the reliability coefficients of two of the types of accidents —collisions with other trolley cars and boarding and alighting accidents—

Table 2

Reliability Coefficients, Coefficients of Correlation, and Distribution Statistics of Collision and Non-Collision Accidents

Trolley Car Motormen. N = 59	Collision Accidents	Non-collision Accidents	Mean	Standard Deviation
Collision accidents	*.42*	.09	7.76	6.04
Non-collision accidents		*.40*	1.14	.89
Motor Coach Operators. N = 34	Collision Accidents	Non-collision Accidents	Mean	Standard Deviation
Collision accidents	*.82*	.13	5.18	2.97
Non-collision accidents		*.80*	1.68	1.40

are quite low, comparisons are difficult. Even with the remaining more reliable measures, however, the intercorrelations do not differ greatly from zero. When the correlations between collision and non-collision accidents, shown in Table 2, are considered, it is again apparent that the relationship between different types of accidents is low.

Conclusions

From the evidence reported by other investigators, together with that presented in the present report, it cannot be considered that accident proneness as a general trait of the individual has been substantiated. Simply because an individual has a high rate for accidents of one kind, it does not necessarily follow that he will have a high rate for accidents of another kind. Rather, if there is any tendency to retain liability to have accidents under many different circumstances, the facts would indicate that such a tendency is of very minor importance as a factor in the determination of accidents.

Received May 10, 1947.

References

1. Newbold, E. M. A contribution to the study of the human factor in the causation of accidents. *Indus. Fat. Res. Bd., Report*, No. 34, 1936.
2. Farmer, E., and Chambers, E. G. A study of personal qualities in accident proneness and proficiency. *Indus. Fat. Res. Bd., Report*, No. 55, 1929.
3. Brown, C. W., and Ghiselli, Edwin E. Factors related to the proficiency of motor coach operators. *J. appl. Psychol.*, 1947, **31**, 477–479.

The Guilford–Zimmerman Aptitude Survey

J. P. Guilford and Wayne S. Zimmerman

University of Southern California

It is the purpose of this article to describe the development of the *Guilford-Zimmerman Aptitude Survey* (7), to cite some of its novel features, and to present some preliminary results.[1]

Reasons for the Survey

The Factorial Approach to Aptitude Testing. First of all, it is our conviction that aptitudes of individuals can be evaluated most adequately, economically, and meaningfully by using a series of tests each of which measures a unique ability (4). Secondly, we believe that the aptitudes required for doing successfully the many kinds of tasks in a complex society are much more numerous and varied than has generally been supposed (2, 6). Thirdly, we emphasize the importance of recognizing more fully than heretofore, that certain of these unique abilities are measured best by "power" tests and others by "speed" tests. Fourthly, the knowledge and understanding of ability factors have grown rapidly during the past few years, especially as a result of war time research, and application should be made of these advances (3).

In the past, tests of intelligence, of clerical ability, and of mechanical ability have served well as far as they go. It is now realized, however, that they do not go far enough. Before the recent war, it had been shown by statistical analysis that human resources, as measured by tests, fall into rather separate and distinct traits to which Thurstone has given the general name of "primary abilities" (9). The accelerated development of knowledge of aptitudes during the war has served to confirm the belief in aptitude factors or primary abilities, and to advance considerably the information concerning them (3).

Advantages of Factor Tests. The most important advantages of a series of tests each of which is relatively unique are (4):

1. *Wide Coverage at Economical Cost.* In most batteries of tests of the traditional type there is so much duplication in the measurement

[1] In this article portions of the test manual (7) written by the same authors have been liberally quoted, without the use of quotation marks, by permission of the copyright holder.

that there are many more tests than there are factors covered. Furthermore, the number of factors actually touched is generally quite limited. Success in a job or assignment often depends upon a large number of factors. For example, it has been estimated that success in pilot training in the Army depends upon at least 20 separate abilities and traits (3). The best way to be sure that all aspects of a job are covered economically, is to recognize the factors and to provide the minimum number of tests needed.

2. *Meaningfulness and Dependability.* When a factor is measured by a test that is unique for that factor, we know rather definitely what the score means. On the other hand, if a single test measures two or more common factors, an individual's score cannot be interpreted with confidence. A better-than-average score in the test may be due to one exceptionally strong ability combined with one or more weak abilities, or it might just as well be due to equivalent strengths in all abilities measured. For use in vocational guidance, tests should yield maximal differences between scores for each person. This is best achieved by means of factor tests. Furthermore, experience has shown conclusively that complex tests do not measure any one factor as well as do factorially "pure" tests.

3. *Adaptation to Batteries.* With a battery of factorially unique tests, one can select a single test for each factor that should be measured, exclude factors that are not needed, and give each test its optimal weight and thus achieve maximal prediction. With complex tests this usually cannot be done, since a good measure of a desired factor may be (and often is) a measure also of an unrequired or undesired factor. Thus, weighting a test weights the undesired factor as well as the desired one.

4. *Enlightened Selection of Batteries.* The best knowledge to have concerning the assignment for which selection is to be made is in terms of the primary abilities required for success in that assignment. Familiarity with the factors and the tests that measure them will often be sufficient to serve as a basis for effective choice of tests until further information becomes available. An inspectional job analysis made with the primary-ability categories in mind is the most effective preliminary approach in this connection. If the job requires rapid and accurate computations as a significant component, then the number factor should be measured. If the job requires careful and accurate inspection of minor details, then the perceptual-speed factor should be measured. The dependability of these better known factors is such that if the observation of the job activities is correct, one would rarely go far astray in the choice of selective tests following the approach recommended.

Plan of the Survey

Traditional Areas Covered. The *Aptitude Survey* has been planned to include, when completed, tests of approximately 20 primary abilities. The seven parts already published cover fairly well the important factors in the three traditional areas—abstract intelligence, clerical aptitude, and mechanical aptitude. The same seven tests may be combined in various ways to measure aptitudes for many other occupational areas.

Parts I and II (Verbal Comprehension and General Reasoning) measure the two leading components of most verbal intelligence tests. Since these two factors are practically independent, it is possible for an individual to have quite different standings in the two tests. For this reason, it is highly desirable to have at least two separate scores to represent this area of aptitude.

Parts III and IV (Numerical Operations and Perceptual Speed) measure leading components of most clerical-aptitude tests. Different kinds of clerical jobs depend to different degrees upon these two primary abilities. Every clerical job must be inspected in its own right for its requirements in terms of human resources. Some clerical jobs also lean more or less heavily upon verbal and reasoning components. Selection for such assignments would call for the use of Parts I and II of the *Survey*. Some jobs might well involve significant amounts of spatial and visualizing abilities which are represented by Parts V and VI. The term "clerical-aptitude" is a very rough occupational concept, not conforming to a single, fixed pattern of psychological functions. This situation calls for the use of different combinations of unique tests, differently weighted to meet special requirements for each type of clerical assignment.

Parts V, VI, and VII (Spatial Orientation, Spatial Visualization, and Mechanical Knowledge), particularly Part VII, cover the most important abilities required in mechanical pursuits. Like clerical aptitude, the concept of "mechanical aptitude" is variable. The common element that justifies the idea of mechanical aptitude is the machine, from which the concept takes its name. But there are many kinds of things one can do with, or in connection with machines. One can invent them, design them, draw blueprints of them, read blueprints of them, construct them, operate them, or repair them. Little reflection is required to realize that the psychological resources important in all these connections are by no means the same. Extensive statistical analyses of mechanical tests have fairly well clarified this area of aptitude (3). The only underlying ability that is unique to mechanical tests proves to be a mechanical-experience factor. This factor seems to depend upon the amount of basic, simple, practical mechanical knowledge that the individual has acquired, formally or informally.

Parts V and VI stress other factors, both of which can be measured independent of the mechanical-experience factor, but which are relatively important in many kinds of mechanical tasks. It is expected that Part V, Spatial Orientation, will be found most useful in prediction of success of street car operators, bus or truck drivers and other machine operators, for the ability measured is found as a component of psychomotor tests in which mechanical devices must be manipulated. Visualization should be most important in connection with designing, and with the making and reading of blueprints, though it is known also to be important in some machine-operating tasks such as piloting an airplane.

New Areas of Aptitude Covered. It should be clear from what has been said that the three traditional aptitude areas, as customarily covered by tests, have been broken down into their chief fundamental aptitude factors and it is believed that those areas can thus be assessed by factor tests in a much more enlightened manner than heretofore. As it was pointed out before, the seven tests go much beyond the three traditional areas. There is no necessity of restricting the groupings of the tests to these three ways. By cutting across these traditional boundaries, a great many other types of vocational and professional aptitudes can be assayed. The various arts and jobs that combine academic and either clerical or mechanical aspects call for different regroupings of tests. When the list of primary abilities tested is later extended beyond the seven, the possibilities for other, non-traditional, groupings will become more apparent.

The Factors and their Measurement. In Parts I through VII, the factors and their measurements may be described as follows (2, 3, 6):

1. *Verbal Comprehension.* This ability is important in any activity that requires understanding or meanings of words or verbal concepts. A general vocabulary test, given as a power test, is the best and purest measure of it. It is commonly evident in such tests as same-opposites, verbal analogies, and reading comprehensions. These three types of tests are likely to measure individual differences in other factors which may not be wanted. Reading-comprehension tests may introduce a variety of factors, depending upon the content of the reading matter used.

2. *General Reasoning.* There are several distinct types of reasoning ability, but the one measured by Part II seems to be common to a larger number of tests than any of the others. It probably makes its strongest contribution in tests composed of items which emphasize the principles characteristic of arithmetic-reasoning problems. From this, and from other facts, it appears that the key to the factor is the ability to diagnose problems. It is best measured in a power test with items of appropriate difficulty.

3. The ability to do rapid and accurate work with numbers stands apart as a distinct human resource. The fact that the extremely rare person may stand out as a "lightning calculator" and yet possibly be generally low in most other abilities is a dramatic demonstration of the independence of this factor. The factor seems to be about equally well measured by any of the four fundamental operations, addition, subtraction, multiplication, and division. Since correct answers to the problems can be derived by nearly all examinees, given sufficient time, the test must be administered under speed conditions.

4. *Perceptual Speed.* This is the ability to perceive detailed visual objects quickly and accurately. Part IV is composed of short matching tests in which the examinee must note similarities and differences in the forms and details of common objects. The items must necessarily be quite easy. If not, they would involve other abilities such as discrimination of size and shape, and perhaps visual acuity and reasoning. Since the items are easy, individual differences must be measured largely in terms of speed.

5. *Spatial Relations.* Part V was designed to measure primarily an ability to appreciate spatial relations of things with reference to the human body. The awareness of whether one object is to the right or left of another, higher or lower, or nearer or farther away, seems to be the essential nature of this factor. Both the proper interpretation of structures and the proper decisions as to how to adjust the body to the layout of machinery depend upon this ability. It was found to be one of the two most important factors for learning to pilot an airplane. It is presumably important in other jobs involving machine operations, since nearly all tests known to contain the factor in some degree appear to have validity for predicting success in mechanical jobs. In Part V, each item shows the prow of a motorboat against a background scene in two similar views. The examinee must report what directions the boat has moved in going from the first to the second of the two pictures. The boat may have turned right or left, may have risen or fallen, and/or may have tilted right or left.

6. *Spatial Visualization.* This factor seems to involve the process of imagining movements, transformations, or other changes in visual objects. It is a dynamic kind of visualization, whereas another factor identified as visual memory is a static or reproductive visualization. It is represented in tests of mechanical movements, mechanical comprehension, and in paper-folding tests of the Binet type. Such tests involve factors in addition to visualization and were therefore not used in the *Survey.* It should be pointed out that before the war the two factors—spatial relations and visualization—had not been separated but were

regarded as one factor called spatial. This distinction has been confirmed by a number of studies (3).

Part VI is composed of pictorial items in each of which a familiar three-dimensional object is first shown in a certain position. Brief verbal instructions then call for turning, rotating and/or tilting the object. The examinee's task is to recognize the object in the new position. Surface-development tests also measure this factor, but nearly always, in our experience, also stress the general-reasoning factor.

7. *Mechanical Experience.* It was indicated earlier that the only factor unique to mechanical tests is an acquired-knowledge or experience factor. Part VII tests the examinee's knowledge of common tools and mechanical problems such as occur around the home, of automobile parts, their functions and malfunctions, and of the common trades of carpentry, plumbing, welding, and the like. It is recognized that not all examinees have had equal opportunity to acquire this kind of knowledge. Girls and women are particularly handicapped in this respect, even though the emphasis is upon more common experiences. It has been found that individual differences in scores made on this test by males are highly prognostic of subsequent success in undertakings of a mechanical nature, e.g., aircraft mechanic and radio operator mechanic.

Construction of the Tests

Some New Features in Test Construction. Several innovations in test development were employed in the construction of the *Survey* tests. In writing items for Part I, Verbal Comprehension, an attempt was made to keep the level of difficulty of all alternative responses equivalent to that of the word to be defined. In some vocabulary tests the responses are somewhat easier than the word to be defined and are of unequal familiarity. It is believed that the discrimination of the right answer from the wrong ones in Part I is more likely than usual to require knowledge of each alternative word meaning. If this is so, each item represents a wider sampling of word knowledge than is customary. Furthermore, the examinee cannot so easily arrive at the right answer by eliminating more familiar (for him) wrong answers. The latter kind of solution, by elimination, is likely to involve reasoning as well as word knowledge. Experience with Part I has shown an unusual tendency for examinees not to guess, as indicated by the large number of omissions.

The approximate difficulty level of each word was determined originally by the use of the Thorndike and Lorge lists (8). Subjective judgment was introduced when it seemed that frequency might not be the only clue to difficulty of meaning. A check upon the success of this method of word selection was made in an item analysis. It was found

that the rank order of difficulty of right responses, as shown by percentages of examinees who chose them, correlated .87 with the difficulty order assumed. In the verbal test the five alternative answers are all of the same part of speech—noun, verb, or adjective—as the word to be defined. Verb forms are clearly indicated by the use of the infinitive form.

In the development of both Parts I and II, longer, preliminary forms (100 items for Part I and 30 items for Part II) were administered to 400 lower-division university students and item analyses were made. The rank order of items for difficulty was established, and each item was correlated with total score in the test. The published forms (Form A of each Part) are composed of items (72 in Part I and 27 in Part II) of wide range of difficulty rather evenly spaced over the range, and of satisfactory correlation with total score. The frequencies of choices of wrong answers to items were studied and some changes were made where it was believed improvements were called for. Item analyses were not made for Parts III, IV, and V because they are speed tests in which the usual item statistics are of little value. Parts VI and VII were not analyzed because on the basis of previous experiences it was believed that the items are functionally homogeneous.

In the construction of Part II, General Reasoning, every effort was made to reduce the amount of impurity with the number factor by reducing the number work to a minimum. Problems are often presented pictorially, in terms of visual units of quantity rather than in terms of numbers. Using a sample of examinees who had taken both the General Reasoning and the Numerical Operations tests, correlation coefficients were computed between each item in the former and total score in the latter. No item was retained in Part II, Form A, that correlated to a significant degree with the number test. The correlation of the preliminary form of Part II with the Numerical Operations test proved to be only .19 in one sample and .20 in another, demonstrating the degree of success achieved before item selection occurred. Part III presents a somewhat novel mode of listing alternative answers, in which the same six alternatives are used to serve two problems in common. This practice conserves upon booklet space and probably upon reading time.

Administrative Features. For all parts, the instructions are made so complete that the tests are largely self-administering. Separate answer sheets can be used with all parts except III and IV.

Some Uses for Which the Tests are Recommended

Kinds of Examinees to Which Adapted. The following list of occupational assignments for which it is believed each part of the *Survey* is predictive, is based partly upon known relatively direct evidence, partly

upon known indirect evidence, and partly upon enlightened guesses. This list is an abridged version of one provided in the test manual (7). It is presented merely for its suggestive value.

Lists of Occupations in Which Factors Measured by the Guilford-Zimmerman Aptitude Survey are Probably Important

Part I. Verbal Comprehension
- author
- editor
- engineer
- historian
- journalist
- lawyer
- navigator
- physician
- scientist
- teacher

Part II. General Reasoning
- accountant
- engineer
- executive
- lawyer
- manager
- mathematician
- navigator
- physician
- scientist
- teacher

Part III. Numerical Operations
- accountant
- actuary
- auditor
- bookkeeper
- calculator
- mathematician
- mathematics teacher
- navigator
- statistical clerk

Part IV. Perceptual Speed
- aircraft pilot
- clerks (most kinds)
- inspectors (of products)
- key punch operator
- navigator
- printer
- proofreader
- typist

Part V. Spatial Orientation
- aircraft pilot
- athlete
- blueprint reader
- dentist
- draftsman
- steam shovel operator
- street car operator
- truck or bus driver
- winch operator

Part VI. Spatial Visualization
- aircraft pilot
- blueprint reader
- dentist
- designer
- draftsman
- electrician
- engineer
- inventor
- machine operator
- navigator
- surgeon

Part VII. Mechanical Knowledge
- aircraft pilot
- carpenter
- chauffeur
- engineer
- lathe operator
- locksmith
- machinist (most kinds)
- mechanic (most kinds)
- repairman (most kinds)
- toolmaker
- tractor operator
- truck driver

Some Preliminary Results

Reliability. For Parts I, II, VI, and VII the odd-even reliability estimates (corrected by the Spearman-Brown formula) are .92, .92, .91,

and .92, respectively. The samples include from about 100 to more than 200 college men. For Parts III, IV, and V, reliability was estimated by administering each test in two separately timed, equivalent halves, intercorrelating the part scores and applying the Spearman-Brown formula. For these three parts the estimates are .92, .92, and .88, respectively (N's = 113, 108, 220).

Validity. There are two types of validity for tests: factorial validity and practical validity (1). Factorial validity is indicated by the correlation of a test with a primary ability. Such a coefficient tells how well the test is measuring a factor. Practical validity is indicated by the correlation of the test with some criterion of adjustment, such as vocational proficiency.

At the time this article was written, there were no practical validity data to report concerning the tests in the *Survey*. Later supplements to the Manual will report such information, since a number of validity studies are in progress.

These particular tests have not yet been factor analyzed, so precise estimates of factorial validity are lacking. It can be stated with some justification, however, on the basis of known factorial composition of very similar tests (3) that the seven tests should be expected to have the following correlations with their dominant factors. Part I, .80; Part II, .60; Part III, .80; Part IV, .75; Part V, .60; Part VI, .60; and Part VII, .80. These factorial validities are probably close to the best obtainable for the factors under present testing conditions.

Intercorrelations. A set of factorially pure tests would have zero intercorrelations if the factors are unrelated. It is believed that the intercorrelations of the seven parts of the *Survey* will be very low, with few exceptions. In these instances a significant intercorrelation may be attributed to the presence of an additional factor or factors common to the two tests. It is possible that in one or two instances it will be desirable to suppress any unwanted factor variance in the test by combining its score in a linear equation with that of another test. Formulas have been provided for such an operation (5).

A few intercorrelations already computed can be reported. Some of these were of special interest because it was anticipated that a few of them would be significantly greater than zero. Between Part I (preliminary form) and Parts II (preliminary form), VI, and VII the correlations are .19, .14, and .10 respectively (N's = 84, 108, 107). It is of interest to note these low figures in the light of the fact that most items in Parts II and VII are presented verbally and that the directions for Part VI probably require greater verbal facility than any of the other tests in the *Survey*.

Between Part II (preliminary form) and Parts III, IV, V (booklet form), and VII the correlations are .19, .13, .27, and .16, respectively (N's = 142, 180, 180, 118). The correlation of .19 is strong evidence that the number factor was reduced to an unprecedented low status in a test made up of items involving the principles found in arithmetic-reasoning problems. The attempt by the authors to construct arithmetic-reasoning-type problems which could be stripped of their numerical content, retaining only the desired reasoning element, was apparently quite successful. This was done on the basis of previous studies which indicated repeatedly that certain types of arithmetic reasoning problems contain the greatest amount of the general-reasoning factor. The low correlation with the perceptual-speed test indicates that the perceptual items were of a sufficiently low level of difficulty that a good general-reasoning ability could not be substituted to solve the problems when perceptual speed was lacking. The low correlation of the reasoning score with the mechanical score is evidence of a proper difficulty level for the latter test and of the omission of items that can be reasoned through to a solution, such as are likely to appear in tests of the mechanical-principles type.

Between Parts III and IV the correlation is .24 (N = 243), between III and V (booklet form) it is .00 (N = 119) and between IV and V (booklet form) it is .28 (N = 116). When separate-answer-sheet recording is used with Part V, the correlations with some other Parts range quite a bit higher. Since the booklet form of Part V obviously is more pure, the use of this form is recommended.

The correlation of the answer-sheet form of Part V with Part VI was found to be .55. When a different scoring formula was tried involving a stronger penalty for errors in Part VI, the correlation was .48. When part scores for Part VI were obtained and correlated with the booklet form of Part V (N = 134), the correlations were found to range as low as .33. No specific recommendations are being made at the present time regarding the application of special scoring formulae or of utilizing part scores. The study is being continued, however, and when more confident interpretation of these correlational variations can be made, the findings will be reported in supplements to the manual.

Norms. Norms are available on a college population and are being accumulated on industrial and high-school populations.

Summary

In the *Guilford-Zimmerman Aptitude Survey* attempts are being made to produce a series of tests each of which features one primary ability. Seven parts already completed cover the most important factors in aca-

demic, clerical, and mechanical and other aptitude areas. New procedures were utilized in the attempt to achieve factorial uniqueness. Preliminary data indicate fairly satisfactory results toward this end. It is proposed that such a battery will serve as a flexible, efficient, and manageable device for assessing individuals in many aptitude areas and will yield unusually interpretable and dependable evaluations.

Received December 1, 1947.
Early publication.

References

1. Guilford, J. P. New standards for test evaluation. *Educ. and Psychol. Meas.*, 1946, 6, 427–438.
2. Guilford, J. P. The discovery of aptitude and achievement variables. *Science*, 1947, 106, 279–282.
3. Guilford, J. P. (Ed.). *Printed classification tests. Army Air Forces Aviation Psychology Program Research Reports, Report No. 5.* Washington, D. C.: Government Printing Office, 1947.
4. Guilford, J. P. Factor analysis in a test development program (to appear in the *Psychol. Rev.*).
5. Guilford, J. P., and Michael, W. B. Approaches to univocal factor scores (to appear in *Psychometrika*).
6. Guilford, J. P., and Zimmerman, Wayne S. Some AAF findings concerning aptitude factors. *Occupations*, 1947, 26, 154-159.
7. Guilford, J. P., and Zimmerman, Wayne S. *The Guilford-Zimmerman aptitude survey.* Beverly Hills: Sheridan Supply Co., 1947.
8. Thorndike, E. L., and Lorge, I. *The teachers word book of 30,000 words.* New York: Teachers College, Columbia University, 1944.
9. Thurstone, L. L. *Primary mental abilities.* Chicago: University of Chicago Press, 1938.

Veterans' Scores on the Purdue Pegboard Test

J. R. Strange and A. Q. Sartain

Southern Methodist University

"The Purdue Pegboard," in the words of its authors, "is a test of manipulative dexterity designed to assist in the selection of employees in industrial jobs requiring manipulative dexterity, such as assembly, packing, operation of certain machines, and other routine manual jobs of an exacting nature."[1] The test consists of a board with two rows of small holes into which metal pegs may be inserted. It yields two types of scores, one for the number of pegs placed in the holes, and one for the number of assemblies (consisting of a peg, a washer, a collar, and another washer) completed. For the placing of pegs scores may be obtained for the right hand alone, for the left hand alone, for both hands working together, and for a total of right hand, left hand, and both hands.

Statement of the Problem

One problem of this study was to determine whether, in the light of the norms furnished by the authors of the Purdue Pegboard, veterans applying for vocational guidance constitute an inferior, an average, or a superior group. Another problem was to intercorrelate the two most important scores yielded by the test and to correlate each of these scores with scores made on the Minnesota Rate of Manipulation Test, the latter yielding a Placing and a Turning score.

Subjects and Conditions of Study

The tests were given at the Southern Methodist University Veterans' Testing Bureau located at Dallas, Texas. This Bureau was set up by the University at the request of the Veterans'Administration for the purpose of vocational testing of veterans. Subjects for this study were picked at random from all the veterans who had taken the test and included both able-bodied and disabled men. Scores obtained from 850 subjects were used to build a set of "local norms" for the group on the Pegboard. An intercorrelation was run between Pegboard Placing [2] and

[1] "Preliminary Manual for the Purdue Pegboard," Purdue Research Foundation, Purdue University, p. 1.

[2] The Placing score referred to is the total for right hand, left hand, and both hands together.

Pegboard Assembly scores, using 447 subjects. For 75 subjects who had taken both the Purdue Pegboard and the Minnesota Rate of Manipulation tests, correlations were run between Purdue Placing and Minnesota Placing, Purdue Placing and Minnesota Turning, Purdue Assembly and Minnesota Placing, and Purdue Assembly and Minnesota Turning.

Results of the Study

Tables 1, 2, and 3 give the results of this study so far as the percentile rank of each score is concerned.[3] Thus, when using the right hand the

Table 1

Percentile Ranks of Scores Made by Male Veterans on Three Subtests of the Purdue Pegboard Test (N = 850)

Score	Percentile Rank		
	Right Hand	Left Hand	Both Hands
23	99.9		
22	99.8		
21	99.1	99.5	
20	97.4	98.9	
19	92.5	96.8	
18	80.9	91.0	99.9
17	61.6	74.8	99.1
16	42.4	55.8	97.3
15	24.9	36.6	90.5
14	12.0	20.5	75.9
13	04.7	09.7	53.1
12	01.3	03.3	28.2
11	00.4	01.4	13.2
10	00.1	00.2	05.1
9	00.1		01.9
8	00.1		00.5
7	00.1		00.1

placing of 21 pegs in the holes in the allotted time was sufficient to put the man in the top 1% of the group. For the left hand the corresponding score was 20, for both hands 17, and for the total 56. Likewise 12½ assemblies were required for ranking in the top 1%. Similar statements could be made about the other percentile ranks.

Table 2 also compares the experimental group with the two groups of subjects on whom the test was standardized. It is evident that our "local norms" for the veterans' group are higher at every point than those for the men on whom the test was standardized. Indeed, the percentile

[3] The reference in all cases is to a single trial and not to the average of three trials.

Table 2

Comparison between Male Veterans' and Standardizing Groups in Total Placing Scores on the Pegboard Test (N = 850)

Score	Percentile Rank		
	Veterans' Group	Standardizing Group	
		Men	Women
62			100
61			
60	99.9		
59	99.8		
58	99.7		
57	99.5		
56	99.1		99
55	98.5		99
54	96.7		98
53	94.6	100	
52	91.5		95
51	87.8		
50	81.5	99	88
49	73.9		
48	66.8	98	78
47	58.9	96	
46	49.1	93	63
45	41.6	87	
44	33.8	80	46
43	27.6	73	
42	21.6	64	29
41	16.4	54	
40	11.3	44	17
39	08.6	34	
38	05.5	26	8
37	03.9	17	
36	02.5	16	3
35	01.8	10	
34	01.1	6	1
33	00.5	3	
32	00.4	2	
31	00.3	1	
30	00.3		
29	00.1		

ranks obtained in the present study are much closer to those obtained from the women, rather than from the men of the standardizing group.[4]

Table 3 gives similar information for the Assembly test. Here the

[4] Since means and standard deviations are not given for the standardizing groups, significance ratios could not be calculated.

Table 3

Comparison between Male Veterans' and Standardizing Groups in Assembly Scores on the Pegboard Test

Score	Percentile Rank		
	Veterans' Group	Standardizing Group	
		Men	Women
14¾	99.9		
14	99.8		100
13¾	99.7		
13½	99.7		
13¼	99.7		
13	99.7		99
12¾	99.5		
12½	99.1		
12¼	98.6		
12	98.1		95
11¾	96.9		
11½	96.2		
11¼	95.4		
11	93.4	100	86
10¾	91.1	99	
10½	87.4	98	
10¼	82.9	97	
10	77.2	96	66
9¾	70.0	94	
9½	62.8	92	
9¼	57.9	89	
9	50.5	86	44
8¾	41.9	75	
8½	35.6	69	
8¼	31.2	65	
8	24.9	57	25
7¾	18.9	45	
7½	15.9	40	
7¼	13.4	33	
7	10.3	28	12
6¾	07.2	19	
6½	05.3	17	
6¼	03.9	14	
6	02.8	12	4
5¾	01.5	8	
5½	01.1	7	
5¼	00.6	5	
5	00.4	3	1
4¾	00.3		
4½	00.2		
4¼	00.1		
4	00.1	1	

same trends are evident, with the women's norms being closer to those obtained in the present study, though even they appear to be somewhat lower.

The final problem of the study was the intercorrelation between Pegboard Placing and Assembly, and the correlations between each of these and the Placing and Turning scores of the Minnesota Rate of Manipulation Test. Table 4 gives the relevant information on this point. The two scores on the Pegboard correlate with each other to about the same extent as the Pegboard Placing score correlates with the two scores of the Minnesota test. As might be expected, the Assembly scores do not correlate as highly with the Minnesota scores. It is evident that these coefficients are relatively low. Whether they indicate low reliability in one or both of the tests or the likelihood that the various tests involve different factors, the present study affords no way of determining.

Table 4

Intercorrelation of Subtests of the Pegboard Test and Correlations of Each with Subtests of the Minnesota Rate of Manipulation Test for Male Veterans

Subtests	*N*	*r*
Pegboard Placing vs. Pegboard Assembly	447	.53
Pegboard Placing vs. Minnesota Placing	75	.50
Pegboard Placing vs. Minnesota Turning	75	.59
Pegboard Assembly vs. Minnesota Placing	75	.33
Pegboard Assembly vs. Minnesota Turning	75	.36

Summary and Conclusions

On the basis of a random sample of 850 male veterans who were seeking vocational guidance and who had taken the Purdue Pegboard test, percentile ranks were calculated for the various subtests, and these were compared with norms furnished by the authors of the test. Using a sample of 447 cases the Assembly and (total) Placing scores on the Pegboard were intercorrelated, and finally with a sample of 75 cases, the Pegboard Assembly and Placing scores were correlated with Placing and Turning scores on the Minnesota Rate of Manipulation test. From the study the following conclusions were drawn:

1. Men's norms as furnished by the authors of the test are definitely lower than those obtained with the group of male veterans.
2. Women's norms as furnished by the authors of the test are considerably closer to those of the male veterans' group.
3. The correlation between the Pegboard Placing scores and Assembly scores was relatively low (.53).

4. Correlations between Pegboard Placing scores on the one hand and Minnesota Placing and Turning scores on the other were also relatively low (.50 and .59 respectively).

5. The Pegboard Assembly scores correlated with the Minnesota scores to a considerably smaller degree (.33 and .36 respectively).

Received March 24, 1947.

High School Norms for the Grove Modification of the Kent–Shakow Formboard Series

Ruth C. Wylie
Connecticut College
Alice W. Wilson
Western Reserve University
and
William R. Grove
University of Pittsburgh

Grove[1] has described a modification of the Kent-Shakow Formboard Series which seems to be fairly homogeneous and which proved difficult enough to discriminate among adults. Wylie[2] has demonstrated that this test is reliable enough at the high school level to warrant further standardization.

Like other investigators in the field of performance testing, Grove found that raw time and move scores yield skewed distributions and non-linear correlations among themselves and with scores from other tests.

Various more or less arbitrary means of weighting raw scores have been used to take care of these difficulties. Grove used a pragmatic weighting table to obtain scores which would yield fairly normal distributions. He fitted free-hand curves as a basis for rectifying his data, after having failed in more precise methods of mathematical curve fitting. The method used was further justified by the fact that the scores thus rectified yielded linear intercorrelations.

The work reported in this paper deals with: (1) the development of a more adequate scoring rationale for use with the test and (2) the construction of high school norms for the test.

Procedure

Administrative Procedure. As part of a larger study, four performance tests and two group tests were administered to 352 boys chosen at

[1] Grove, W. R. Modification of the Kent-Shakow Formboard Series. *J. Psychol.*, 1939, 7, 385–397.

[2] Wylie, R. C. The reliability of the Grove modification of the Kent-Shakow Formboard Series. *J. appl. Psychol.*, 1947, 31, 155–159.

random from grades 7–12 of the Beaver Falls, Pennsylvania, Public Schools. The following tests were administered:

1. Grove Modification of the Kent-Shakow Formboard Series.
2. Block Designs Test from the Wechsler-Bellevue Scale.
3. Cube Construction Test from the Cornell-Coxe Performance Scale.
4. O'Connor's Wiggly Block Test.[3]
5. California Test of Mental Maturity, Short Form.
6. Revised Minnesota Paper Formboard Test, Series AA.

Characteristics of the Normative Group. Since the scores of this group of subjects were being used to develop norms for the Modified Kent-Shakow Formboard Series, it was necessary to inquire first into the general characteristics of this sample.

Beaver Falls is an industrial community of approximately 18,000 people, not far from Pittsburgh, Pennsylvania. The town is predominantly middle class with a generous proportion of pupils of foreign extraction in the High School.

Table 1

Means and Standard Deviations of Rate (Recesses per Minute) Scores on the Modified Kent-Shakow Formboard Series (352 Boys in the Junior-Senior High Schools of Beaver Falls, Pennsylvania)

		Subtest							
		A		B		C		D	
Grade	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
7	54	8.44	2.83	2.17	1.22	2.41	1.26	1.10	0.57
8	56	8.52	2.59	2.54	1.50	1.89	1.23	1.04	0.66
9	62	9.68	2.46	2.96	1.44	2.48	1.06	1.29	0.53
10	54	9.71	2.56	3.30	1.58	2.70	1.34	1.15	0.75
11	63	9.80	2.58	3.53	1.51	2.90	1.09	1.20	0.71
12	63	10.29	3.12	3.30	1.38	2.94	1.28	1.32	0.67
9–12	242	9.88	2.71	3.27	1.49	2.74	1.19	1.24	0.67
7–8	110	8.49	2.71	2.36	1.38	2.10	1.27	1.08	0.61

The mean chronological ages progressed from 12 to 18 in accord with what one might expect in a random sample from these grades.

Tables 1, 2, 3, and 4 summarize the data obtained from the tests administered to this group. The mental ages obtained from the California Test of Mental Maturity tend to run slightly above those expected for these grades. The agreement is close between the Block Design Scores made by this group and Wechsler's norms for this test. On the Revised

[3] Later eliminated from consideration due to its unreliability.

Table 2

Arithmetic Means and Standard Deviations of Mental Ages* of the Subjects Used in Studying the Grove Modification of the Kent-Shakow Formboard Series at the High School Level

Grade	N†	M.A. Language Factors		M.A. Non-language Factors		M.A. Total Mental Factors	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
7	53	12.8	1.5	13.2	1.6	13.1	1.7
8	52	14.0	1.6	13.5	1.4	14.3	1.4
9	49	16.0	1.6	15.6	3.2	15.7	1.6
10	53	16.8	1.8	16.8	3.7	16.8	1.8
11	52	16.8	2.3	17.1	3.2	16.8	2.2
12	54	17.1	1.9	16.7	3.4	16.8	1.8

* California Test of Mental Maturity Intermediate S-Form (Grades 7–8) and Advanced S-Form (Grades 9–12).

† N = 313 in this table, this being the total number of cases for whom California Test of Mental Maturity Scores were available. As far as is known no systematic selective factor operated in determining which of the 352 subjects missed the group testing sessions.

Minnesota Paper Formboard our group tends to run a little higher than the Likert and Quasha norms, the greatest difference appearing in the group whose chronological ages run from 16–25. This difference may be due to the fact that our 16–19 year-old group was not strictly comparable in age range or school status to Likert and Quasha's 16–25 year-old group.

Table 3

Arithmetic Means and Standard Deviations on the Wechsler Block Designs for the Sample Used in this Study as Compared with the Norms Used by Wechsler

C.A.	Present Study			Wechsler		
	Mean	S.D.	N*	Mean	S.D.	N
11	8.1	1.9	8	7.4	2.4	60
12	9.8	2.6	33	9.1	2.5	60
13	9.5	2.8	35	10.1	2.9	70
14	10.4	2.8	55	10.7	2.9	70
15	10.8	3.0	52	10.7	3.2	100
16	11.2	2.5	64	10.9	3.2	100
17–19	10.9	3.2	60	11.0	2.9	100

* N = 307 in this table, this being the number of cases for whom the following data were available: California Test of Mental Maturity Scores, chronological ages (between 11 years and 19 years), and Wechsler Block Designs Scores.

On the whole, the sample used in this study seems to be rather similar to various other samples which have been used in previous standardizations of these supplementary tests.

Table 4

Comparison of Medians and Quartiles on the Minnesota Paper Formboard, Revised Series AA, for the Sample Used in This Study with the Norms Given by Likert and Quasha*

Groups	Present Study				Likert and Quasha			
	Median	Q_1	Q_3	N	Median	Q_1	Q_3	N
C.A. 16–25 (males)	41	35	47	116	34	26	39	147
C.A. 14–6 to 15–5 (males)	40	30	48	46	38	31	44	101
C.A. 11–6 to 12–5 (males)	30	25	34	21	32	26	38	96
High school seniors	42	35	47	53	39	33	45	1288

* The groups included are the only ones from Likert and Quasha's norms which were comparable to the subjects of the present study.

Use of the critical ratio and analysis of variance techniques showed that the grades 9–12 constituted a sufficiently homogeneous group with respect to the formboard to justify combining their formboard scores into a "pool." This pool was then used as the standardization group in constructing the weighting tables and the norms. (See Tables 5 and 6.)

Table 5

Summary of the Data on Variance for the Modified Kent-Shakow Formboard Series

(Between the means of the recesses per minute scores of grades 9, 10, 11 and 12, critical points are: $F_{.05} = 2.65$ $F_{.01} = 3.88$)

Subtests	F
A	1.52
B	1.50
C	1.58
D	1.22

Table 6

Summary of Critical Ratios for the Differences between the Mean Recesses per Minute Scores on the Modified Kent-Shakow Formboard Series

Subtest	Grades		
	9–12 vs. 7–8	9–12 vs. 8	7 vs. 8
A	4.48	3.48	0.15
B	5.69	3.32	1.43
C	4.57	4.47	2.17
D	2.28	2.00	0.50

Results

Scoring Rationale. Although both time and moves were recorded for each subtest for each individual in the standardization group, we decided, for several reasons, to discard moves as a scoring element. First, Wylie (*loc. cit.*) had demonstrated that scores based on moves were less reliable than scores based on time, and that moves combined with time did not improve the reliability of the test. Second, the correlation between rectified time and moves scores was +.84. Finally, moves are more difficult to count and record than time. For these reasons, it was decided to base the scoring upon time alone.

As noted earlier, Grove (*loc. cit.*) had demonstrated that raw time scores were unsatisfactory. It was decided to try rate scores. The basic scoring unit adopted was the number of recesses correctly completed per minute. For each subtest rates were calculated by dividing the number or recesses correctly completed by the number of minutes taken to complete them. Thus, since each subtest consisted of five recesses, a subject taking five minutes to complete any given subtest would have a rate score for that subtest of one recess per minute. Similarly, if a subtest was completed in one minute, the rate score would be five recesses per minute; while if it were completed in thirty seconds, the rate score would be ten recesses per minute.

The use of these rate scores proved to be a satisfactory solution to the difficulties encountered with raw time scores. First, it was found that rate scores bore a substantially linear relationship to the pragmatic weights used by Grove to rectify his raw time scores. At each extreme of the table, Grove's weights deviated slightly from a linear relationship with rate scores, but through the middle range the linearity was satisfactory. Thus the rate scores were functionally comparable to Grove's pragmatic weights through the major portion of his scoring table. Second, the method of rate scores yielded more symmetrical distributions

for all subtests than the distributions obtained from raw time scores. Third, linear intercorrelations between the various subtests and between the subtests and other tests resulted from the use of rate scores. In contrast, the same regressions were non-linear when raw time scores were used. Finally, the use of rate scores provides a satisfactory rationale for scoring partially completed subtests on a basis similar to completed subtests. Thus, rate scores for partial success (completion of one to four recesses within the time limit for any subtest) can be computed in terms of the same basic unit as for the fully completed subtests. This method obviates the necessity for assigning arbitrary or pragmatically weighted point values to partial successes as previous performance test investigators have done.

Since the Modified Kent-Shakow Formboard Series consists of separate subtests, some method of weighting the subtests was necessary. If the subtests were permitted to weight themselves upon the basis of the rate scores obtained for each, undue weight would be given to the easier subtests at the expense of the more difficult.

Ideally the weights assigned to the subtests should be determined on the basis of a validity study, i.e., against an external criterion. However, a validity study was beyond the scope of the present work. Wilson[4] had tried assigning weights to each subtest in direct proportion to its relative difficulty. Cases in the pool 9–12, scored according to this method, yielded a distribution which was a poor fit for the normal curve. Other possible methods were statistically explored. We were impressed by the very high intercorrelations obtained between these various methods of weighting the subtests. We accordingly decided that the simplest and most direct method was to let each subtest contribute equally to the total score.

It may seem to the reader that such a scoring procedure would be cumbersome and time consuming in practical use. This would indeed be the case except for the fact that the entire process has been reduced to the use of a single conversion table. (See Table 7.) This table lists the weighted scores to be assigned for each subtest, depending upon the number of seconds taken by the subject on completed subtests, or the number of recesses complete at the time limit, in the case of unfinished subtests. If no recesses were completed at the time limit, the score for that subtest was zero. The scoring table has the further merit that it has been statistically adjusted in such a way as to yield total weighted scores for our normative population (Grades 9–12) which have a mean of 50 and a standard deviation of 10.

[4] Wilson, A. W. Educational norms for high school boys on the Modified Kent-Shakow Formboard Series. Unpublished Ph.D. Thesis, University of Pittsburgh Library, 1944.

Table 7

Weighted Scores Corresponding to Raw Time Scores for the Modified Kent-Shakow Formboard Series (242 High School Boys, Grades 9–12)*

Weighted Score	Subtest A	Subtest B	Subtest C	Subtest D
1	—			
2	I			
3	II			
4	III and IV			
5	90–120			
6	71–89	I	I	—
7	59–70	II, III, IV	II	—
8	50–58	255–360	III and IV	I and II
9	43–49	178–254	203–240	III and IV
10	38–42	138–177	159–202	401–480
11	34–37	112–137	132–158	307–400
12	31–33	94–111	112–131	249–306
13	29–30	82–93	97–111	209–248
14	26–28	72–81	86–96	181–208
15	24–25	64–71	77–85	159–180
16	23	58–63	70–76	142–158
17	21–22	53–57	64–69	129–141
18	20	49–52	59–63	117–128
19	19	45–48	55–58	108–116
20	18	42–44	51–54	100–107
21	17	39–41	48–50	93–99
22	16	37–38	45–47	87–92
23	—	35–36	43–44	81–86
24	15	33–34	40–42	77–80
25	14	31–32	38–39	73–76
26	—	30	37	69–72
27	13	28–29	35–36	65–68
28	—	27	33–34	62–64
29	12	26	32	59–61
30	—	25	31	57–58

* These scores have been statistically adjusted to give total scores which are standard scores with a mean of 50 and an S.D. of 10.1 for our normative group.

The steps that in Table 7 are so conveniently telescoped into one simple operation, are illustrated by the example shown in Table 8. This example is based on the raw scores of a single subject of our normative group. Column two gives the raw scores made by the subject. In

Table 8

Scoring Example

Subtest	Raw Score	Rate Score Recesses per Minute	Sigma Deviation of Subjects' Rate Score from Mean	Weighted Score from Col. I, Table 3
A	48 sec.	6.25	−1.34	9
B	114 sec.	2.63	−0.43	11
C	200 sec.	1.50	−1.04	10
D	3 recesses at time limit	0.38	−1.28	9
			Total Weighted Score =	39

column three, these raw scores are converted into rate scores. To illustrate how these rate scores were computed we may note that the five recesses of Subtest C were completed by this subject in 200 seconds or 3.33 minutes. Thus, the subject's rate score for Subtest C was 1.5 recesses per minute (5/3.33). On Subtest D the time limit is eight minutes, and in that time the subject completed three recesses. Thus, in

Table 9

Distribution of Total Weighted Scores Obtained on the Modified Kent-Shakow Formboard Series by the Standardization Group (242 High School Boys)

Total Score	Frequency
75–79	3
70–74	2
65–69	10
60–64	20
55–59	41
50–54	52
45–49	53
40–44	32
35–39	17
30–34	7
25–29	0
20–24	2
15–19	1
10–14	1
5–9	1
N	242
Mean	50.0
S.D.	10.1

Subtest D, he was working at the rate of 0.38 recess per minute (3/8). Column four lists the sigma deviation of each of this individual's rate scores from the mean of the normative group. These values were obtained by use of Table 1. If these sigma deviation values were algebraically totalled we would have a score which would give equal weight to each subtest. In column five the same result is more expeditiously accomplished by using the scoring table (Table 7) in which equal weight is given to each subtest with the added advantage that *the total weighted score is actually a standard score.* Thus, the individual in our example obtained a total weighted score of 39 which is 1.1 sigmas below the mean of high school boys. By using Table 7, one can thus avoid doing such

Table 10

Percentile Scores for the Standardization Group Corresponding to Total Weighted Scores on the Modified Kent-Shakow Formboard (242 High School Boys)

Percentile	Total Score	Percentile	Total Score
99	74	50	50
98	70	40	48
97	68	30	45
96	67	20	42
95	66	10	38
90	62	5	34
80	58	4	33
70	55	3	30
60	53	2	27
50	50	1	20

computations as are illustrated in columns three and four of the example, yet achieve comparable results.

Norms for High School Boys. As has already been explained, our has scoring table (Table 7) has been so constructed as to incorporate the norms for a homogeneous and representative group of high school boys.

The actual distribution of total weighted scores obtained for our group is presented in Table 9. This distribution is a good fit for the normal curve (chi square test). Table 10 has been constructed to give the percentile values corresponding to the standard scores.

Discussion

Practical experience with the formboard and correlations between the formboard and other tests, support the hypothesis that this test measures an aspect of general intelligence, but will probably be found

more heavily weighted with the spatial visualization factor (Wylie, *loc. cit.*). Validity studies are now badly needed to increase further the usefulness of the series.

Although this test requires bulky and expensive equipment, and must be individually administered, it has some very good features which should commend it to the clinician and counselor:

1. It has intrinsic interest for adult subjects;
2. It is self-corrective;
3. It has an apparent validity which commends it to the mechanic or engineer who may doubt the efficiency of pencil and paper tests;
4. It does not penalize the individual who lacks verbal skills;
5. It can be administered with relative ease to a completely deaf person;
6. It is apparently able to discriminate adequately over a wide range of ability;
7. As performance tests of this type go, it is a reliable instrument;
8. It takes a relatively short time for administration and scoring (approximately fifteen minutes).

The test is not now commercially available, but a manual is being prepared and it is hoped that a reliable test publisher can be interested in manufacturing and distributing the test.

Received March 22, 1947.

Visual Skills and Labor Turnover

N. C. Kephart

Division of Education and Applied Psychology, Purdue University

A relationship has been repeatedly shown between visual skills of workers and such industrial problems as production efficiency, quality of work, training time, safety, etc. A further difficulty in industrial personnel procedure is the problem of labor turnover. Much needless expense is inccurred in the training and breaking in of new employees who after such training remain on the job only a short time. Industrialists are therefore interested in determining the factors which cause new employees to quit after a short period on the job and as far as possible eliminating these factors so that employees can be hired with a fair assurance that they will stay on the job until their work has become profitable both to themselves and to the company. The present paper will show how, in a specific company, vision was found to be one of the factors contributing to short tenure of workers.

A large manufacturer of optical goods noted a decidedly high rate of turnover in his lens inspection department. The jobs in this department involved the inspection of eye glass lenses for surface quality. After cleaning the lens, the operator looks for physical defects such as chips, scratches, holes, bubbles and so on. This work is highly repetitive but requires more than the average visual attention since many of the defects to be noticed are small and difficult to detect visually.

As a part of the employment procedure in this company, each applicant for employment is given the battery of visual tests incorporated in the Ortho-Rater before he is hired.[1]

Through this pre-employment testing program, visual test scores were available for all workers who had been placed on this job irrespective of the length of time they remained on the job. From the records a group of 32 workers who had remained on this inspection job for a period of 8 months or more was identified as a sample of workers remaining on the job. Similarly a group of 67 workers who had been placed on this job and who remained less than 4 months was identified as a sample of workers who left the job shortly after they were hired. These two groups

[1] The Ortho-Rater is a visual testing device manufactured by the Bausch & Lomb Optical Co. for the visual classification and placement of industrial employees.

were chosen so as to represent reasonably extreme ends of the tenure distribution in order to determine whether available personnel and test data would differentiate between employees who did not stay on the job even long enough to complete the training and those who remained sufficiently long to justify their initial employment. By reference to the preemployment test file, visual test scores were obtained for all workers in each of these two groups. The visual test scores of the workers who remained more than 8 months were compared with the scores of those

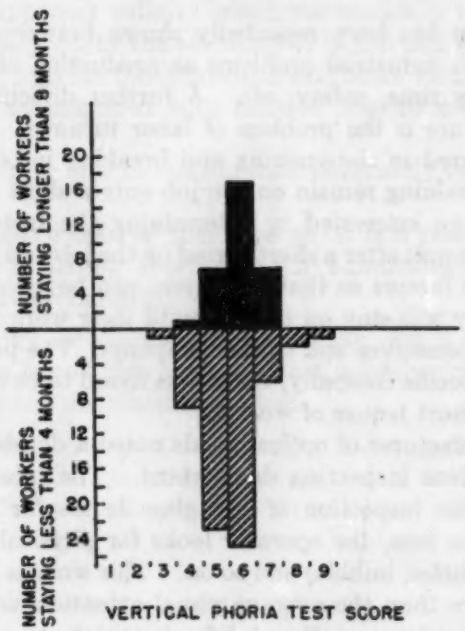


FIG. 1. Distribution of scores on test of vertical phoria.

workers who remained less than 4 months to determine the relationship between visual skill and length of tenure on this job.[2]

Figure 1 shows the distribution of scores on the test of vertical phoria at the optical equivalent of 26 feet. Under certain test conditions, which eliminate the necessity for the eyes to converge on a single point, the eyes assume a posture that may converge or diverge from that required in normal seeing at the test distance. Such postures, measured in terms of angular deviation from that required normally for that distance, are considered the phoria condition. In the battery of tests used in this

[2] The method of analysis used in this comparison has been described previously. Tiffin, Joseph, and Wirt, S. E. Determining visual standards for industrial jobs by statistical methods. *Trans. Amer. Acad. Ophthal. and Otolaryn.*, 1945, 50, 72–93.

study such postures are measured in both a vertical direction and lateral direction. There are, therefore, two such phoria tests, a vertical phoria and a lateral phoria.

From the distribution in Figure 1 it can be seen that the number of individuals scoring at either extreme on this test is greater among those workers who remain less than 4 months than among those who remain more than 8 months. Only 3 individuals scored 8 or 9. None of these 3 individuals remained on the job more than 4 months. Only 11 individuals scored less than 5. Of these, 10 remained on the job less than 4 months and only 1 more than 8 months.

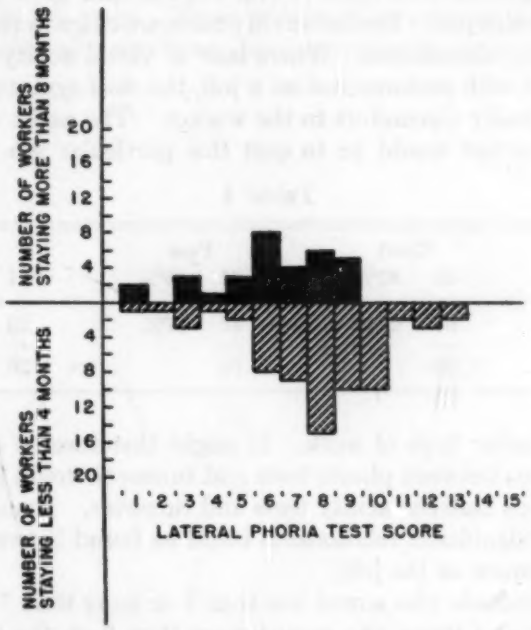


FIG. 2. Distribution of scores on test of lateral phoria.

The standard deviation of scores for those individuals remaining more than 8 months was .75. The standard deviation of scores for those individuals remaining less than 4 months was 1.07. This is a difference of .32 which is 2.46 times its standard error. It thus appears that workers who remain less than 4 months are more variable in vertical phoria than are workers who remain more than 8 months. This conclusion can be drawn at the 2% level of confidence.

Figure 2 shows the distribution of scores for the two groups on the test of lateral phoria at the optical equivalant of 26 feet. It will be seen that 17 individuals show a score greater than 9 on this test. Of these 17

individuals none remained on the job more than 8 months. The mean score of workers remaining on the job more than 8 months was 6.22. The mean score of workers remaining on the job less than 4 months was 7.93. This is a difference of 1.71 which is 3.43 times its standard error. The difference between these means is significant at the .1% level.

It would therefore appear that those workers who show a deviation in vertical phoria in either direction and those workers who show high scores in lateral phoria (towards the exophoric end of the scale) more chacteristically remain on the job less than 4 months.[3]

Since this lens inspection job requires prolonged visual concentration it is not unexpected that these phoria tests should show a relationship with tenure on this job. Deviations in phoria are characteristically associated with visual discomfort. Where lack of visual acuity characteristically interferes with performance on a job, the chief symptom of a phoria is characteristically discomfort to the worker. The easist way to alleviate such discomfort would be to quit this particular job or to request

Table 1

	Total		Pass		Fail	
Stay	32	32%	31	43%	1	4%
Quit	67	67%	42	57%	25	96%
	99		73		26	

transfer to another type of work. It might therefore be expected that the relationships between phoria tests and turnover would be closer than the relationships between acuity tests and turnover. Among this group of workers no significant relationship could be found between the visual acuities and tenure on the job.

Those individuals who scored less than 5 or more than 7 on the vertical phoria test and those who scored more than 9 on the lateral phoria test were identified and their length of tenure compared with that of those workers whose scores did not fall in these undesirable ranges. The resulting figures are shown in Table 1.

It will be seen that 43% of those individuals whose test scores fell

[3] In general, distance acuity tests show greater correlation with success on distance jobs and near acuity tests show greater correlation with success on near jobs. Distance acuity tests have shown low relationships with near acuity tests (Tiffin, Joseph. Vision and industrial production. *Illuminating Engineering*, XL, No. 4 (April 1945)). Similar relationships, however, have not in general been shown for the distance and near phoria tests. The present finding, therefore, that distance phoria tests are related to turnover on a "near" job is not inconsistent with these earlier findings dealing primarily with acuity tests.

within the desirable ranges remained on the job more than 8 months while 57% remained on the job less than 8 months. On the other hand, among those individuals whose test scores fell within the undesirable ranges, only 4% remained on the job more than 8 months while 96% remained on the job less than 6 months.

This relationship was also investigated by the computation of the tetrachoric *r*. This value was .71.

These data would indicate that for visual inspection jobs such as the one studied here, there is a definite relationship between visual phorias and job tenure. Those individuals who deviate from ortho-phoria in either direction vertically and toward exophoria laterally are much more apt to leave the job early. These relationships between the phorias and tenure are statistically significant both individually and in combination.

Received May 24, 1947.

The Validity and Reliability of Heterophoria Scores Yielded by Three Commercial Optical Devices *

J. H. Sulzman, M.D.
Troy, New York
Lt. Comdr. E. B. Cook, USNR
Naval Medical Research Laboratory, New London, Conn.
and
N. R. Bartlett
Johns Hopkins University

An earlier communication (1) discussed the relaibility of visual acuity measurements with the Bausch and Lomb Ortho-Rater, the American Optical Sight Screener, and the Keystone Telebinocular. That report also presented data on the agreement between visual acuity scores with these devices and scores obtained with a standard acuity test. The evaluation of these three instruments is extended in this note to the measurement of heterophoria. The same 121 observers and the same conditions of testing were used in this second phase of the study as in the first, since the data were recorded simultaneously. The testing procedure is sketched in its essentials in the previous article (1) and is described in detail in the original Navy research reports (2, 3) from which these communications are abstracted. To summarize briefly, each person of a representative population of observers was examined twice with each of the three instruments and twice with a standard clinical test, with the order of testing randomized. Then for each method under investigation the test-retest reliability coefficient, the distribution of scores for the initial test and for the second test, and the coefficient of correlation of the test scores with the criterion standard measure were computed. This evaluation is based entirely on these statistical indices.

The Criterion Clinical Test

Several tests for heterophoria are used by ophthalmologists in clinical practice. Three used widely are the Maddox rod, the Screen-Maddox

* This investigation was carried out at the Medical Research Laboratory, Submarine Base, New London, Conn. The opinions expressed herein are the private views of the authors, and are not to be construed as official or as representing the Naval service at large.

rod, and the Screen and Parallax tests. The three appeared to have about the same face validity. A preliminary experiment with one hundred subjects was conducted therefore to settle upon one of the three as a criterion. The experiment showed the three to be highly intercorrelated, so a choice on the basis of test-retest reliability, brevity and simplicity was indicated. The measure chosen was the Maddox rod test. Statistical data for this test are presented later in Tables 1 and 2.

In the Maddox rod test, a subject sees a small spot of light with his left eye, and with his right eye sees the same light as a vertical streak displaced to one side. Ordinarily the subject does not realize that actually he is seeing each object monocularly. He is shown that the line of light may be displaced laterally by rotating a Risley prism, and then he is asked to adjust the prism until he is satisfied that the line passes through the center of the spot. The rotation is made rapidly by most observers without pausing or reversing the direction of movement, and the first setting for each observer is reported ordinarily to be in good coincidence. The score for the test is based on the lateral deviation (in prism diopters) introduced in achieving coincidence. For the vertical measure of imbalance, the Stevens phorometer attachment is used instead of the Risley prism and the Maddox rod is rotated to present a horizontal streak instead of a vertical streak.

Maddox rod tests for lateral and vertical phoria were conducted in this study with the far target at twenty feet, and the near target at thirteen inches. The former was at the same level as the eyes of the observer, and the latter six inches below eye level. This depression approximates the eye-muscle conditions characteristic of ordinary reading. Altogether, four phoria measures are yielded by this Maddox rod procedure; a vertical and a lateral imbalance score when the eyes are focussed on the target at distance, and a vertical and a lateral score when they are converged on the target close at hand.

Table 1 presents means and standard deviations of the four scores for the initial and for the second tests of the sample of 121 observers. Results are expressed in conventional terms of prism diopters. Latent deviations from optical alignment in the lateral plane are called exophoria (X) if the test indicates the eyes have a tendency to diverge, and esophoria (E) if they tend to converge. On the other hand, latent deviations in the vertical plane are designated according to the eye which tends to deviate above the other; thus, RH denotes hyperphoria for the right eye, and LH hyperphoria for the left.

No large systematic changes in score distributions from the initial to a second test are revealed in Table 1. The extent to which such changes occur, and the degree to which relative scores for a population are

Table 1

Test and Retest Means and Standard Deviations for Maddox Rod Test

	Means		Standard Deviations	
	Test	Retest	Test	Retest
Far Distance				
Lateral	E 1.30	E 1.12	3.78	3.89
Vertical	LH 0.08	RH 0.28	0.54	0.39
Proximate Distance				
Lateral	X 3.84	X 3.04	6.02	5.65
Vertical	RH 0.50	RH 0.50	0.61	0.57

consistent from one test to another as indicated by test-retest correlation data, together afford an evaluation of reliability.

Reliability and intercorrelation coefficients for the four tests are shown in Table 2. Those in parentheses are test-retest coefficients. The latter are high, but of course are not high enough to warrant much reliance on a single administration of the test for individual prediction or diagnosis. On the other hand, whatever their value, these reliability coefficients can serve as a yardstick for assessing corresponding data for the three instruments. The small magnitude of the remaining coefficients in the table, indicating the relationship of the several tests to each other, signifies relative independence of the four separate indices of imbalance.

Commercial Devices Under Comparison

As stated in the previous section, the Maddox rod test displays a spot of light to the left eye and a streak of light to the right eye. The same general principle of presenting dissimilar targets to the two eyes is employed in all three of the instruments studied.

While the testing procedure with each of the three is essentially the

Table 2

Reliability and Intercorrelation Coefficients for Maddox Rod Tests

	Far Distance		Proximate Distance	
	Lateral	Vertical	Lateral	Vertical
Far Distance				
Lateral	(0.793)			
Vertical	0.151	(0.623)		
Proximate Distance				
Lateral	0.668	0.155	(0.867)	
Vertical	0.173	0.446	0.159	(0.677)

same, the Sight-Screener differs in one major respect from the other two. With all three devices, two dissimilar objects are seen, one with each eye; one object appears stationary, while the other may appear to drift. The drifting object seen by one eye is made in the form of an indicator, which appears to point to some one of a row series of dots or steps seen with the other eye. The subject reads the value at which the free indicator finally comes to rest in the method prescribed for both the Telebinocular and the Ortho-Rater. In the Sight-Screener procedure, however, the extreme excursion to which the indicator swings is read instead of the position at which it comes to rest.

In addition to this procedure difference, there are a few mechanical differences in the devices. Telebinocular targets are illuminated from the front, while the Sight Screener and Ortho-Rater targets are trans-illuminated from behind. Decentered convex lenses are used for viewing in every case except with the Sight Screener for proximate distance; in the latter instance, targets are viewed directly without the interposition of prismatic lenses. In addition, the Sight Screener is unique among the three in that the Polaroid Vectograph principle is employed for presenting separate targets to each eye. Finally, at the time of this investigation, there was no Telebinocular test available for near vertical phoria.

Statistical Data and Evaluation

Test and retest means, standard deviations and test-retest reliability coefficients are presented in Table 3. Score units are expressed in the raw scale units of the respective instruments.

On the basis of comparative reliabilities, as estimated from the stability of test and retest means, from the constancy of test and retest standard deviations and from test-retest reliability coefficients, there seems to be no clear-cut argument for choosing any one of the instruments in preference to any other.

The degree of relationship between the first test and the first corresponding Maddox rod test is presented for each instrument in Table 4. For purposes of this comparison, that relationship is an index of validity. The validity coefficients, test-retest reliability coefficients, and the validity coefficients corrected for attenuation are shown.

The raw validity coefficients in Table 4 (column 2) are low, and furthermore are considerably smaller than the reliability coefficients. So the validity indices after correction for attenuation are still well below 1.00. Evidently, then, the instruments do not tap the same psychophysiological functions as do the corresponding four Maddox rod tests. No hypotheses are offered to account for the fact that the same functions are not involved; many possible reasons might be proposed to explain it,

but the fact itself is the important thing. But before any interpretations about validity are made, the criterion should be examined. It must be emphasized that the choice of the Maddox rod test for this purpose was arbitrary. Probably the Screen-Maddox rod or the Screen and Parallax tests would yield validity coefficients similar to the Maddox rod test, for the three are closely interrelated. But the principal argument for employing any one of the three rests on widespread clinical usage. The validity figures might have been different had some other kind of criterion been used.

Table 3

Test and Retest Means, Standard Deviations and Reliability Coefficients for the Three Devices

	Means		Standard Deviations		Reliability Coefficient
	Test	Retest	Test	Retest	
Ortho-Rater					
Far Lateral	6.79	6.59	2.72	2.56	0.872
Far Vertical	5.17	5.11	1.07	1.01	0.630
Near Lateral	8.08	7.51	2.76	2.69	0.924
Near Vertical	4.50	4.59	0.83	0.83	0.624
Sight Screener					
Far Lateral	14.94	14.60	1.48	1.92	0.796
Far Vertical	3.06	3.02	0.37	0.29	0.613
Near Lateral	17.07	16.30	3.33	2.97	0.831
Near Vertical	3.02	3.06	0.34	0.39	0.551
Telebinocular					
Far Lateral	E 1.48	E 2.20	2.31	3.58	0.755
Far Vertical	4.71	4.73	0.84	0.76	0.602
Near Lateral	E 0.48	E 1.16	5.69	5.46	0.880

In general, the correlation between instrument measures for lateral imbalance and corresponding Maddox rod scores is greater than for vertical measures. Furthermore, the reliability coefficients reveal a greater consistency for tests of lateral phorias than for tests of vertical phorias. Of course, this finding may be attributed to the greater relative homogeneity of the population with respect to vertical imbalances; for example, the standard error in predicting retest scores is less for vertical phorias than for lateral. However, if standard score scales were to be constructed separately for lateral phorias and for vertical phorias on the basis of this population, then the test-retest consistency in terms of these standard scores would be much greater for lateral measures than for vertical.

Table 4 affords a direct comparison of test-retest coefficients of the clinical tests with the coefficients for the several corresponding instru-

ments. This comparison reveals no loss in reliability in substituting the machine for the clinician. The only question to be settled before recommending the use of the more expeditious and less troublesome mechanical test in place of the clinical procedure is that of comparative validities. So long as one holds that the mechanical devices must measure exactly the same functions as does the conventional Maddox rod test, then the devices cannot be recommended without reservations. Perhaps a more reasonable statement is that the machines when used with the manufacturer's recommended procedure present tests differing

Table 4

Reliability Coefficients of the Clinical Measure and the Instrument Measure and the Intercorrelations between the Two

Name of Test	Coefficient of Reliability of Clinical Measure	Coefficient of Interrelation between Instrument and Clinical Measure	Coefficient of Reliability of Instrument Measure	Intercorrelation Coefficient Corrected for Attenuation
Ortho-Rater				
Far Lateral	0.793	0.564	0.872	.68
Far Vertical	0.623	0.286	0.630	.46
Near Lateral	0.867	0.674	0.924	.75
Near Vertical	0.677	0.343	0.624	.53
Sight Screener				
Far Lateral	0.793	0.370	0.796	.47
Far Vertical	0.623	0.279	0.613	.45
Near Lateral	0.867	0.543	0.831	.64
Near Vertical	0.677	0.337	0.551	.55
Telebinocular				
Far Lateral	0.793	0.371	0.755	.48
Far Vertical	0.623	0.426	0.602	.70
Near Lateral	0.867	0.683	0.880	.78

in some unknown fashion from the conventional clinical procedure, but that on the other hand they are at least as reliable as the clinical method.

Conclusions

Perhaps the most useful conclusion is that this comparative evaluation of heterophoria measurements with the three instruments does not indicate any clear-cut basis for preferring any one instrument to the other two. Secondly, no one of the three offers sufficient reliability to warrant the use of a fine scale for lateral or vertical phoria when the test is administered only once. Thirdly, the machines offer reliabilities at least as great as certain clinical tests now accepted and in use. Finally, the

question of validity is not settled, for the scores with the machines do not correlate to a satisfying degree with scores on one standard clinical test.

Received April 3, 1947.

References

1. Sulzman, J. H., Cook, E. B., and Bartlett, N. R. Reliability and validity data for certain commercial devices for measuring visual acuity. *J. appl. Psychol.* (In Press.)
2. Sulzman, J. H., Cook, E. B., and Bartlett, N. R. Visual acuity measurements with three commercial screening devices. Bur. Med. Surg. Research Project No. X-493 (Av-263-p), Progress Report No. 2; 7 Feb. 1946.
3. Sulzman, J. H., Cook, E. B., and Bartlett, N. R. Comparative study of measures of heterophoria. Bur. Med. Surg. Research Project No. X-493 (Av-263-p), Progress Report No. 3; 22 Feb. 1946.

Construction and Use of Weighted Check-List Rating Scales for Two Industrial Situations *

Edwin B. Knauft

State University of Iowa

Several previous investigators (1, 2, 3,) have used the Thurstone equal-appearing intervals method to select and weight items for use in a check-list type of personnel rating device. However, in each of these instances the obtained items were then subjected to an extensive item analysis procedure which served as the basis for the final form of the rating device. Since such item analyses generally require an independent criterion measure of the ratees' job performance and involve a considerable expenditure of time and labor, the use of a simplified technique should extend the use of this type of rating method to a broader range of industrial situations. The present paper describes the relatively simple procedures involved in the construction of rating scales of the check-list type by the use of the equal-appearing intervals method. These techniques have been applied to the construction of rating devices for the two very different jobs of laundry press operator and bake shop manager.

Construction of a Scale for Laundry Press Operators

The rating scale for laundry press operators[1] was originally constructed in connection with a validation study of a selection test for these operators. The problem was somewhat unique because there is an average of only eight press operators employed in the typical laundry. Since satisfactory test validation was impossible in any one laundry due to the small number of cases, it was necessary to obtain data on operators from a number of different laundries under different managements. It was not possible to obtain satisfactory objective measures of the production of each operator, and thus it was necessary to resort to a merit rating method as the sole criterion of operator proficiency. The final rating device took the form of a weighted check list of items which were selected and weighted by the Thurstone equal-appearing intervals

* Originally presented in part at the 1947 meetings of the Midwestern Psychological Association.

[1] The writer is indebted to Mr. H. H. Hauser for assistance in the collection of data on the final form of the press operator rating scale.

method. The following steps were involved in the construction of the scale:

1. Each of 25 laundry managers was asked to prepare a list of statements which described the performance of a press operator. Each statement was to express in a simple sentence one thing an operator did which would aid or hinder her production or the production of other workers around her. About 600 such statements were collected, but after editing and eliminating the duplicates, 197 statements were selected.

2. The 197 statements were submitted to 27 laundry managers for sorting. Each statement was reproduced on a separate card and the managers were instructed to sort the statements into nine piles which represented nine points on a good-to-poor press operator continuum. Thus each statement was evaluated in terms of the value to the laundry management of the described behavior.

3. The median scale value of each statement was computed on the basis of the position assigned to the statement by each of the 27 sorting managers. The semi-interquartile range of each statement was similarly computed.

4. All statements with a Q value of 1 scale unit or more were discarded, leaving 150 statements for further analysis.

5. The final scale was constructed from the remaining statements on the basis of the median scale value of each statement. Twenty-seven statements with scale values at approximate intervals of 0.3 scale unit were thus selected for the final scale. The statements with the smallest Q values were selected wherever possible.

6. A second form of the scale was constructed which contained 27 other items having approximately the same values as the statements in the first form.

7. The items in each form were then placed in random order and the ratee was instructed to mark each statement as being true or not true of the operator being rated. The score for the individual ratee is obtained by computing the mean of the scale values of statements which were checked as being true or descriptive of that ratee. The scale values of the statements are not shown on the rating check list and are not known by the raters.

The final items included in Forms A and B of the Press Operator Rating Scale are listed in Tables 1 and 2. These tables also give the scale or scoring value of each item. It will be noted that several items appear in both forms of the scale. This procedure was necessary because several intervals along the continuum were each represented by only one item of acceptable Q value. In such instances the same item had to be included in both forms of the scale.

Table 1

Items of Form A—Press Operator Rating Scale

Item	Scale Value
She efficiently utilizes all her press area in making lays.	8.3
She shows moderate interest in her work.	4.6
She is not very careful with her equipment.	2.8
She is fairly systematic in her work.	5.8
She works at a slow but steady pace.	4.3
She is absent from work frequently.	1.1
It is hard for her to change her ways of doing things.	4.1
She can change to a different work without much loss of speed.	7.3
She uses too many lays.	3.0
She makes few errors in lay sequence.	6.9
It upsets her to change work.	3.8
She does not give enough straight pull to the garment before closing the press.	6.5
Her work seldom requires any extra touch-up.	7.9
She is fair at breaking in new workers.	5.7
She is careful to keep the garments clean.	7.0
She pats and smooths the lays more than necessary.	3.6
She seldom asks for any time off.	7.7
She tries to pass the blame on to others.	2.0
She is more interested in quality than speed.	6.4
She can't seem to get along with anyone.	0.6
She picks over items in the damp box too much.	3.3
She can turn out good work at high speed.	8.4
She frequently threatens to quit.	0.7
She accepts constructive criticism well.	8.1
She does a sloppy job of pressing.	0.9
She never seems to know when to use the damp cloth.	2.4
She is careful of her equipment.	7.9

Results Obtained With Press Operator Rating Scale

The rating scale was used by 15 different laundry managers to rate their press operators. Data were thus obtained on 118 ratees who were rated on each of the two forms of the scale. The reliability coefficient of the scale, based on the scores obtained on the two forms by the 118 operators, was .87. The reliability coefficient for the full scale of both forms, as estimated by the Spearman-Brown prophecy formula, was .93.

A composite rating score for each of these operators was obtained by computing a mean score, based on the score the operator obtained on both forms. This distribution of scores had a mean of 5.91 and a standard deviation of 0.81. These scores were subjected to an analysis of variance to determine if there were significant differences between the ratings made by the 15 raters. The resulting F (of 1.02) is not significant

Table 2

Items of Form B—Press Operator Rating Scale

Item	Scale Value
She has the idea the work is "beneath her."	0.8
She does a thorough job of pressing.	8.2
She is a trouble maker.	0.5
She misses quite a bit of work because of illness.	1.9
She gives every evidence of being interested in her work.	7.4
You can never predict how she will take criticism.	3.3
When she speeds up, her quality falls off.	4.3
She uses poor judgment in the sequence of garments.	2.8
She is fair at breaking in new workers.	5.7
She doesn't utilize every opportunity to use double lay.	3.9
She makes some unnecessary movements in handling garments.	4.3
She can satisfactorily operate any kind of press unit.	8.1
She is an extremely slow worker.	1.4
She literally drops the garment into proper lay on the press.	8.4
She wants to be waited on.	1.8
She is seldom distracted from her work.	7.0
She is more interested in quality than speed.	6.4
She is indifferent about the quality of her work.	2.5
She uses good judgment in the use of the damp cloth.	7.5
Her work requires more than normal touch-up.	2.5
She always has a chip on her shoulder.	1.0
She is occasionally late to work.	4.3
She is fairly systematic in her work.	5.8
She makes few errors in lay sequences.	6.9
She does not give enough straight pull to the garment before closing the press.	6.5
She does not load her presses to efficient capacity.	3.6
She has to be pushed continually.	1.6

at the 5 per cent level of confidence, indicating that the differences between mean ratings made by the several raters may be attributed to chance.

It should be noted that a number of the ratings were made by raters who were located at distant laundries where personal contact and instruction in rating procedures was impossible. These raters, who had no previous experience with formal rating methods, made their ratings on the basis of written instructions sent to them by mail. This experience serves to demonstrate that extensive training of raters *may* not be a requisite with this type of rating method.

A Rating Scale for Retail Bake Shop Managers

A second application of the weighted check-list type of rating scale has been made in a study of bake shop managers employed by a large

bakery chain. Each of these managers has responsibility for the operation of a retail bake shop in which all types of baked goods are produced and sold to the public. The items were collected and scaled by the same procedures as were used with the press operator scale with the one exception that the statements for the manager scale were placed into nine categories by a modified and more simple technique. This simpler method requires each judge to evaluate each statement by encircling one of the numbers from 1 to 9 following the statement. The numbers refer to positions on a continuum which ranges from "descriptive of a manager of no value to the company" to "descriptive of a manager of very great value to the company." A study by Seashore and Hevner (4) has indicated that such a procedure will produce statement scale values which are not significantly different from the values obtained by the more cumbersome card sorting procedure.

One hundred and ninety-six statements were scaled by 17 judges who were either district managers, assistant district managers or home

Table 3

Items of Form A—Bake Shop Manager Rating Scale

Item	Scale Value
He occasionally buys some of his competitor's products.	6.8
He never consults with his head salesgirl when making out a bake order.	1.4
He belongs to a local merchants' association.	4.9
He criticizes his employees unnecessarily.	0.8
The window display is usually just fair.	3.1
He enjoys contacting customers personally.	7.4
He does not know how to figure costs of products.	0.6
He lacks a long range viewpoint.	3.5
His products are of uniformly high quality.	8.5
He expects too much of his employees.	2.2
His weekly and monthly reports are sometimes inaccurate.	4.2
He does not always give enough thought to his bake orders.	1.6
He occasionally runs a selling contest among his salesgirls.	6.8
Baking in his shop continues until 2 P.M. or later.	8.2
He keeps complaining about employees but doesn't remedy the situation.	0.9
He has originated one or more workable new formulas.	6.4
He sometimes has an unreasonably large inventory of certain items.	3.3
Employees enjoy working for him.	7.6
He does not delegate enough responsibility to others.	2.8
He has accurately figured the costs of most of his products.	7.8
He wishes he were just a baker.	0.8
His shop is about average in cleanliness.	4.4
He is tardy in making minor repairs in his sales room.	1.9
He periodically samples all of his products for quality.	8.1

Table 4

Items of Form B—Bake Shop Manager Rating Scale

Item	Scale Value
His window display always has customer appeal.	8.5
He gives his employees the reasons for his decisions.	6.7
Products dropped on the floor are sometimes sold.	1.4
He always gets his reports in on time.	7.9
He rarely figures the costs of his products.	1.0
He belongs to a local merchants' association.	4.9
His bakers tend to pass some of their work off onto him.	3.1
He seldom forgets what he has once been told.	7.6
He occasionally runs a selling contest among his salesgirls.	6.8
He is slow to discipline his employees even when he should.	1.9
He does not anticipate probable emergencies.	2.4
His weekly and monthly reports are sometimes inaccurate.	4.2
He should take more interest in merchandising.	3.5
He is slow at making decisions.	4.4
His bakers do not respect him.	0.8
His sales per customer are relatively high.	7.4
He has originated one or more workable new formulas.	6.4
No baking is done in his shop after 12 noon.	0.6
He encourages his employees to show initiative.	8.1
He knows how but can't teach others.	2.5
His shop is unusually neat and clean.	8.3
He often has vermin and insects in his shop.	0.8
His salesgirls sometimes fail to use wax paper to handle goods in the sales room.	3.2
He plays little attention to his customers.	1.6

office officials thoroughly familiar with actual bake shop operations. Two equivalent forms of a check-list scale were constructed by methods similar to those used with the press operator scale. Each form contained 24 items. Tables 3 and 4 contain the items comprising the two forms of the Bake Shop Manager Rating Scale.

This scale was used to evaluate 79 bake shop managers. The basic data were obtained from the ratings resulting when each district manager applied both forms of the scale to his shop managers. The reliability of the scale, based on the product moment correlation of scores on the two forms, was .79. The reliability of the combined scale of both forms was estimated by the Spearman-Brown formula to be .88. Additional data were available on 35 of the above managers who were also rated by their respective assistant district managers. The reliability coefficient of the scale, based upon the ratings of the thirty-five managers evaluated by two raters, was .81.

An estimate was also made of the relationship between the rating

scores and ranks obtained by asking each district manager to rank his store managers in order from best to worst. Although the resulting rank orders cannot be considered as entirely independent measures of the managers' on-the-job-ability, a comparison of the rankings with the scale scores obtained by these ratees does furnish a rough validity estimate of the scale. A rank difference correlation was computed between the rating scores and the rankings each district manager made of his managers. The average of these correlations, based on the ratings and rankings of 79 ratees by seven district managers, was .59. If the correlation of one rater (*rho* = .02) is omitted, the mean *rho* becomes .69 for the remaining six raters who represent 71 ratees.

A composite score for each of the 79 ratees was computed on the basis of the rating resulting when each district manager applied both forms of the check-list scale to the store managers in his respective district. The mean of this distribution of scores was 5.51, with a standard deviation of 0.87 and a range from 3.51 to 7.08. An analysis of variance of these scores was made to determine if there were significant differences in ratings made by different raters. The resulting F value of 2.15 is not significant at the 5 per cent level of confidence, indicating that purely chance differences probably explain the differences in mean ratings between the seven raters.

Discussion

The relatively simple techniques here utilized have yielded scales which compare favorably in reliability with those constructed and refined by more tedious and precise item analysis methods. The equal-appearing intervals techinque as applied to merit rating scale construction has thus far received very limited application in industry in spite of several advantages of this type of scale over conventional graphic or linear scales. These advantages include:

1. The application of a systematic method of utilizing the opinions of a large number of "experts" in the selection and weighting of the items, rather than relying on the judgment of one scale constructor.
2. The ease with which two equivalent forms of the same scale may be constructed, enabling the investigator to obtain a precise reliability measure of the instrument.
3. An objective and rapid method of scoring the completed forms.
4. The exact values or weights of each item are unknown to the raters and may not be readily deduced by them.
5. Requires very few directions to the raters and no extensive training program in rating techniques.
6. May be used as an over-all criterion measure because it may be

constructed to sample a large number of different aspects of the employee's on-the-job behavior.

7. In specific industrial situations this type of scale has exhibited satisfactory reliability and has yielded distributions of acceptable form and range.

Summary

1. The Thurstone equal-appearing intervals technique has been used to select and weight items for two check-list merit rating scales which were used to evaluate personnel. Unlike previous investigations of a similar nature, the present procedure does not make use of an independent criterion in the construction of the final scale.

2. Two forms of a rating scale for laundry press operators were constructed. Managers of 15 laundries rated 118 press operators on both forms of this scale. The reliability of the total scale of 54 statements as determined from the two forms was .93.

3. In a similar manner a rating scale was constructed for managers of bake shops in a large chain. The reliability of this total scale, based on two forms, was .88 (N = 114). The ratings of pairs of raters correlated .81.

4. The rating scores obtained on the manager scale were subjected to an analysis of variance and the resulting F value indicated that purely chance differences probably explain the differences in mean ratings between the several raters. Similar results were also obtained in the case of the press operator scale.

Received May 22, 1947.

References

1. Ferguson, L. W. The development of an adequate method of appraisal. *Amer. Psychol.*, 1946, **1**, 279 (Abstract).
2. Ferguson, L. W. The development of a method of appraisal for assistant managers. *J. appl. Psychol.*, 1947, **31**, 306–311.
3. Richardson, M. W., and Kuder, G. F. Making a rating scale that measures. *Person. J.*, 1933, **12**, 36–40.
4. Seashore, R. H., and Hevner, K. A time-saving device for the construction of attitude scales. *J. soc. Psychol.*, 1933, **4**, 366–372.

Communication Between Management and Workers *

Donald G. Paterson and James J. Jenkins
University of Minnesota

The general problem of communication between management and workers has been receiving increased emphasis in recent industrial personnel literature.

History of the Problem

Roethlisberger (**19**), Roethlisberger and Dickson (**20**), Gardner (**7**), Mayo (**12**), and Pigors and Myers (**15**) may be taken as examples of the recent emphasis on the factory as a social institution in which problems of structure, status, coordination and cooperation of groups and communication between top management and rank and file workers are stressed and discussed at length. For example, there are 21 references to communication in the index of Gardner's book. But in these approaches no hint is given as to how written communications should be formulated or how the problem of readability can be attacked in an objective manner. In Pigors and Myers (**15**) the employee handbook is recommended as the principal medium of communication with stress on content. The only reference to readability is that it be "clearly written" but no information is given as to how this can be accomplished or, if attempted, how the readability of the content may be measured.

In the more comprehensive texts on industrial relations and personnel work such authors as Yoder (**26**), Scott, et al. (**21**), and Watkins and Dodd (**25**) discuss the problem of communication between management and workers in a more practical manner in terms of the makeup and content of employee handbooks, house organs, employee magazines, and bulletin board notices. But here again, although the admonition may be given that these should be written and edited for clarity and interest-holding power, no information is given as to how this can be done or as to how readability may be determined. Watkins and Dodd do give one

* The writers have used the conventional term management and workers throughout this paper. Dr. Dale Yoder, Director of the Industrial Relations Center at the University of Minnesota, suggests that a better terminology would be "managements and employees" because management is an abstraction. Furthermore, he believes that the term employees is to be preferred to workers since the latter includes all members of the labor force (managers, self-employed, rank and file employees, and others).

hint, namely, that some employee magazines have failed because they have been "too high brow" for the man in the shop.

Heron (8a) devotes a whole book to the problem of sharing information with employees and, in our opinion, provides an excellent discussion of the compelling reasons for the necessity of such sharing as a prime function in industrial relations work. Unfortunately, the technical problem of how-to-do-it, in the sense of insuring comprehensibility and interest-holding power of printed communications, is not touched upon.

Filipetti, in his survey (3) of the literature of scientific management from the days of Frederick W. Taylor through World War II, does not mention the problem of communication. The index to his book does not contain any items such as communication, employee handbooks, bulletin board notices or readability. Presumably, information about scientific management procedures and techniques is transmitted from management to workers by verbal means only.

Even Powell and Schild (18), writing for the American Management Association and attempting to tell management how to prepare and publish an employee manual, fail to come to grips with the problem of readability. Their manual summarizes the results of a survey of company practices and gives selected samples from "good handbooks." It also states, "Try to make it attractive and as readable as possible" (18, p. 22) but the description of a technique for testing the readability of the language of employee handbooks is conspicuously absent.

The industrial psychologists do no better. Oakley (13), Tiffin (24), Maier (11), Burtt (2), and Poffenberger (17) do not even mention the problem of communication nor such media as employee handbooks, house organs, or bulletin board notices.[1]

The strange thing is that applied psychologists in dealing with the psychology of advertising devote a great deal of space to the problem of making advertising copy readable. See Kitson (10), Poffenberger (16) and Burtt (1). Not only is the problem recognized by these authors but sufficient facts and descriptions of methodology are given to aid the advertising copy writer in making his copy readily comprehensible to the "average man." As a matter of historic fact, Kitson (10) as early as 1921 developed an objective method of measuring the differences in read-

[1] Recently, increasing attention has been given, however, by industrial relations experts to the problem of simplifying the communication of financial statements to employees and by these same experts and personnel psychologists to the problem of arranging adequate means of communicating the feelings and attitudes of the rank and file to top management through suggestion systems and attitude and morale surveys. In opinion polling a great deal of attention has been given to the wording of questions so that they will be well within the comprehension level of the "average man" (21, 22, 24, 25, 26).

ability or comprehensibility and applied it to newspaper and magazine copy. He used frequency with which one, two, three, etc. syllable words are found, coupled with the average length of sentences. He seems to have been the real pioneer.

The above paradox is most striking in that Burtt and Poffenberger, when they shift from advertising to vocational, employment and industrial psychology, fail to stress or rather fail to mention the problem of communication and hence give no hints as to how to measure the readability of copy in communication from management to the workers. Poffenberger (17) in his book on applied psychology devotes only 37 pages to advertising yet stresses readability whereas he devotes 227 pages to vocational, employment, and industrial psychology without mentioning the problem of communication in the latter situations. Burtt in his 568 pages on employment psychology (2) likewise fails to mention the problem although he devotes adequate space to the problem of readability in his book on advertising.[2]

The problem of readability, however, has been treated far more adequately by Smith, Lasswell, and Casey (22), by Gray and Leary (8) and more recently by Flesch (4, 5, 6). Again, as a matter of historic fact, W. S. Gray seems to have been alert to the need for an index of readability. As early as 1935 he had developed mathematical formulae for measuring degrees of readability and applied them in obtaining indexes of difficulty of 350 books. Unfortunately his method does not seem to have "caught on."

Of course, at a still earlier date Thorndike (23) and Horn (9) had produced word counts that were used not only by authors in writing text books for children but also by advertisers in simplifying the vocabulary of advertising copy.

All of this historical introduction points to a "cultural lag" [3] in applying to personnel work what has long been known in the field of advertising and in the writing of text books for children and for the "average adult." To be sure, the problem of communication between management and workers may have been regarded as of minor importance. But in-

[2] The senior writer pleads guilty to the same offense because when he taught advertising psychology and personnel psychology simultaneously during the 1920's he stressed readability in advertising but failed to recognize its applicability in personnel work.

[3] The phenomenon of "cultural lag" described so ably by Ogburn (14) is important and ever present in any field of knowledge or activity. In psychology many instances could be cited. In this particular matter, the lag seems to have been prolonged by compartmentalized thinking in which people as consumers are recognized as of primary importance whereas people as workers are accorded a secondary role from the point of view of their needs for an understanding of the policies and practices followed by top management.

dustrial relations experts and personnel psychologists can no longer ignore the problem especially in view of the importance attached to communication by those who are now stressing the social structure and human relations approach to management problems in business and industry. The effectiveness with which the Flesch formula can now be applied to the problem also forces the issue.

A Sample Study

In setting up a problem dealing with the selection of power sewing machine operators in a textile factory, the writers were impressed by the attempt of the personnel manager to put in the hands of applicants a printed statement in regard to employment opportunities. Here was a time saver. No longer need the preliminary interviewer give a verbal description of the Company and its hiring policies prior to application taking. Here was a conscious attempt to give each potential applicant a picture of the situation so that each could decide whether or not to put in an application. The applicant information sheet (Form A) is reproduced as Figure 1.

A glance at the wording suggested to the writers that it was probably pitched at a readability level far beyond the comprehension capacity of the average applicant for factory work. This led to an attempt to simplify the language and structure and to increase its human interest value without changing the ideas contained in Form A. This was done by following the rules of Flesch (6). The resulting Form B is shown as Figure 2.

APPLICANTS ** GENERAL INFORMATION—FORM A

Before completing your application it would be advantageous for each applicant to know something about our Company, the products we manufacture, and the employment opportunities for both experienced and inexperienced persons.

TYPE OF WORK—Our Company manufactures a very high quality line of women's and children's underwear, women's nightwear and slips.

IS IT NECESSARY TO HAVE EXPERIENCE—Many organizations follow the policy of employing only experienced operators. Although it is desirable that applicants have had some training, our records indicate that many of our highest-paid operators have had no previous experience or training before they entered our employ. Sincere interest in this type of work is much more important than previous experience in some other organization.

CAN ANYONE LEARN THIS WORK—We do not encourage everyone who makes application to enter this industry. We recognize that some applicants are not interested or suited for this work, just as there are women who are not suited by temperament to enter office work, nursing, or some similar vocation. It would be unfair to you as an applicant, and to the Company if we were to adopt a policy of hiring all applicants.

HOW LONG DOES IT TAKE TO BECOME A GOOD OPERATOR • A power sewing machine, like a typewriter, comptometer or calculator, is an intricate piece of equipment. A student entering business college spends six months acquiring the fundamentals of typing. Speed and proficiency are acquired only with practice and experience in a business office. The average person with an aptitude for sewing can learn to operate a machine in from 2 to 3 months and become an experienced and skilled operator in from 4 to 6 months. We are required to make an expenditure of more than $300 for the training for each operator.

HOW MUCH CAN I EARN AS A POWER SEWING MACHINE OPERATOR—All persons entering our employ are guaranteed 60¢ an hour for the first 3 months and 65¢ an hour thereafter. Our plant average for persons on piece work and having a minimum of 6 months' experience or more is approximately $1.00 an hour. We do not wish you to assume that you will be paid $1.00 an hour at the end of your training period or at the end of the 6 months' period, however, fair piece work rates assure you the opportunity of earning in proportion to your ability. We desire to train only those persons interested in continuous employment.

WHAT ARE THE ADVANTAGES OVER OTHER TYPES OF WORK OPEN TO WOMEN—Persons employed in this industry and by this organization are assured of permanent employment, at light, clean, interesting work which most women enjoy. Many industries have openings for only unskilled or semi-skilled work, such as food packaging, stock work, and retail selling. These positions, for the most part, require very little training and the earning opportunities are therefore definitely limited.

IS IT POSSIBLE TO DETERMINE WHETHER AN APPLICANT CAN BECOME A SUCCESSFUL OPERATOR—It is difficult because of the great differences in individuals to be 100% accurate in selecting potentially successful operators. We know that most persons who have normal vision and above average finger dexterity will be successful if they are truly interested, have patience, and will carefully follow the instructions of our training supervisors.

If you are interested in this employment opportunity, it will be necessary for you to complete the attached Application and Supplementary Information Schedule. You will then be given an eye test in this office and referred to the U. S. Employment Service where you will take several short simple tests. Our records indicate that, all other factors being equal, persons who do well in these tests usually become skillful highly-paid operators.

We wish to state in conclusion that if you are selected for employment we will do everything possible to assure your success. Our instructors will assist and teach you during your training period; and they will be constantly available for guidance and direction. During the past year many persons have entered our employ and have become a permanent and satisfied part of this organization.

FIG. 1. Copy of original information sheet for potential applicants (Form A).[4]

[4] The reader will notice that the section headings are set in *all capitals* in Form A, as in the original, but that *caps and lower case, bold face* are used in Form B. This has been done in order to direct attention to the importance of using the best possible typographical arrangement to promote speed and ease of reading the copy. This is another important though frequently neglected aspect of the problem of printed communications. See Paterson, D. G. and Tinker, M. A. *How to make type readable.* New York: Harper and Brothers, 1940. Pp. 209. (Obtainable from the writers.)

Information for Applicants—Form B

Before You Apply. We think you would like to know something about our Company. This page tells you what we make and how *you* will fit into the jobs we offer.

What Do We Make? We make a very high quality type of women's and children's underwear, women's nightwear and slips. We make these with electric sewing machines like those in many homes.

Do You Need Experience? Some companies only hire people who have done this kind of work before. Some companies demand experience. We do not. Why? Because our records show that many of our best-paid workers started work here without experience. Some had never worked before. Now they are very good workers. We like to hire trained workers but we know that real interest in sewing and real ability come first.

Can Anyone Learn This Work? No. Some people are not interested in it or suited for it. You know some women who are not good bets for office work, nursing or other special jobs. In the same way, some women are not good bets for our jobs. It would not be fair to you if we hired you for a job you did not like or could not do. If we hire you, it means we think you are really fitted for this work.

How Soon Can You Become a Good Worker? People in business school spend six months learning to type or run an electric calculator. Speed and good work come only with practice in school and on the job. Most people we hire can learn to sew well on our electric machines in 2 or 3 months. Some can learn in less time. It usually takes 4 to 6 months to become highly skilled. It costs us more than $300 to train you. This is why we have to be careful about hiring the right kind of worker.

How Much Can You Earn in our Company? You are *guaranteed* 60¢ an hour for the first 3 months and 65¢ an hour after that. Most of our workers with 6 months' experience are making $1.02 an hour. Fair piece work rates give you the chance to earn according to your ability. We like to hire only people who want steady work. And we want those people to be able to earn $1.00 an hour or more.

What Are Your Advantages Here?

1. People hired by us are sure of *steady* work.
2. Most women enjoy this work which is light, clean and interesting.
3. Pay is higher for this work than for unskilled or partly skilled work such as stock work, food wrapping or selling, which most firms in Minneapolis offer. Such jobs can be learned in a week or two and pay is therefore much less than $1.00 an hour.
4. The women who work here are the kind you will enjoy working with.

How Can You Find Out if You Can Be Good in this Work? This is hard to do because there are great differences between people. But we know that most people who have normal eye-sight and nimble fingers will succeed. Of course, you must be interested in sewing machine work, have patience and follow the training of our teachers.

Can You Get Work With Us? Yes, if you will do these things and measure up to our standards:

1. Fill in the Application Blank.
2. Answer Supplementary Information Schedule.
3. Take an eye-sight test in this office.
4. Go to the U. S. Employment Office, 501 1st Avenue North, where you will take several short simple tests for nimble fingers. Our records show that those who do well in these tests usually become good operators with high earnings.

Will We Help You Succeed? Yes. If you are hired we will do our best to help you make good. Our teachers will train you and help you during your training period. When you are ready for regular work, our foremen and workers will help you to make good. All of us want you to make good and to become a satisfied worker.

FIG. 2. Copy of simplified information sheet for potential applicants (Form B).

A glance at Form B is enough to reveal that it is couched in much simpler language (shorter sentences, easier vocabulary content) and may be expected to have a higher human interest value (use of pronouns "you" and "we") than is true of Form A.

The question, however, arises as to precisely how difficult Form A is and to what degree the difficulty of Form B is decreased as compared with Form A. Furthermore, how much better adjusted is the language of Form B to the comprehension level of the "average applicant"?

To answer these questions we employed the Flesch method of measuring readability to all sections of both forms. The results are shown in Table 1. In applying the Flesch formula (6) units of 100 words each

Table 1

Flesch Readability Measurements and Difficulty Interpretations of Information Sheets for Applicants, Forms A and B

Form A	Readability Score	Difficulty Level	Typical Magazine	School Grade Level
Section 1	6.49	Very hard	Scientific	College
Section 2	5.77	Hard	Academic	H.S. or some college
Section 3	5.46	Hard	Academic	H.S. or some college
Section 4	5.41	Hard	Academic	H.S. or some college
Section 5	6.77	Very hard	Scientific	College
Section 6	4.89	Fairly hard	Quality	Some H.S.
Total (all sections)	5.81	Hard	Academic	H.S. or some college
Form B				
Section 1	1.96	Easy	Pulp mag.	5th grade
Section 2	1.68	Easy	Pulp mag.	5th grade
Section 3	2.27	Fairly easy	Slick fiction	6th grade
Section 4	2.02	Fairly easy	Slick fiction	6th grade
Section 5	2.27	Fairly easy	Slick fiction	6th grade
Section 6	0.96	Very easy	Comics	4th grade
Total (all sections)	1.84	Easy	Pulp mag.	5th grade

were measured as sections. Section 6, since Form B was shorter, did not run to 100 words so extrapolation was resorted to.

It will be noted that Form A had been written at the "Hard to Read" level of difficulty appropriate for high school or college students. There is some variability from section to section from "Fairly Hard" to "Very Hard." It should be noted that the first section of Form A was written at the level characteristic of scientific writing requiring college level attainment to understand.

Form B on the other hand was written on an "Easy" level which can be comprehended by those with only a fifth grade educational attainment level. There is also some variability from section to section ranging from "Very Easy" to "Fairly Easy."

The objection may be raised immediately by some that by presenting such simple 4th, 5th and 6th grade level language that the personnel department runs a grave risk of insulting the intelligence of the applicants most of whom have probably had some high school training. One need only re-read Form B to realize that simplification of language need not result in "childish" writing nor is there anything in the information as given or as it is stated that would insult the intelligence of anyone. Bright people can quickly understand Form B. But dull people or even "average" people cannot understand or at least will have great difficulty in understanding text material suitable for college students. While we believe that Form A deserves the label "high brow," we do not believe that Form B deserves the label "low brow." If this be accepted, there can be no quarrel with the simplicity of Form B.

One other point deserves mention. Form B includes many more personal pronouns which Flesch insists is a legitimate technique for enhancing the interest value of any type of writing whether it be "Very Hard" or "Very Easy."

Summary

1. The literature of industrial relations and personnel work has either ignored the problem of communication between management and workers or has failed to emphasize the importance of communicating in language that the "average man" can understand and has given no hints as to the techniques of readability measurement that can be applied.

2. The literature of advertising, however, is in marked contrast. Here the necessity of putting advertising copy into simple language has been emphasized and ways and means of doing so have been stressed.

3. Although Kitson in 1921 and W. S. Gray in 1935 developed ways and means of testing the readability of printed prose, it remained for Flesch to develop and "put across" a tedious but accurate method of

readability measurement. His method can now be applied to communications between management and workers as well as in many other fields.

4. A sample study was made of an information sheet for potential applicants in a needle trades factory. Its Flesch readability score classified it as "Hard," typical of an academic type magazine and requiring a high school or some college level of reading ability to comprehend it. A simpler form was prepared, preserving the ideas, and the resulting Flesch readability score was found to be "Easy," typical of pulp fiction writing and comprehensible to those with a 5th grade level of reading ability.

5. The rules set forth by Kitson (**10**), Thorndike (**23**), Horn (**9**), and Flesch (**6**) for making printed copy easy to understand and at the same time interesting should be followed by those who wish to prepare copy that will be readily understood by *applicants for factory jobs.* These are: (*a*) Use short sentences; (*b*) use words of one or two syllables; (*c*) use simple sentence structure; (*d*) use words in frequent use by consulting word lists; (*e*) avoid unnecessary adjectives and (*f*) make the message personal by using personal pronouns.

6. The literature on industrial relations and personnel psychology should, henceforth, not only place greater emphasis on the importance of communication but should devote much more attention to the "how-to-do-it" aspect of insuring communications that will be readable to the man in the shop and will rate high in interest value.

Received November 29, 1947.
Early publication.

References

1. Burtt, H. E. *The psychology of advertising.* New York: Houghton Mifflin Co., 1938. Pp. 473.
2. Burtt, H. E. *Principles of employment psychology.* Revised Edition. New York: Harper and Brothers, 1942. Pp. 568.
3. Filipetti, G. *Industrial management in transition.* Chicago: Richard D. Irwin, Inc., 1946. Pp. 311.
4. Flesch, R. Estimating the comprehension difficulty of magazine articles. *J. gen. Psychol.*, 1943, 28, 63–80.
5. Flesch, R. *Marks of readable style.* New York: Columbia University, Teachers College, Contributions to education, No. 897, 1943. Pp. 69.
6. Flesch, R. *The art of plain talk.* New York: Harper and Brothers, 1946. Pp. 210.
7. Gardner, B. B. *Human relations in industry.* Chicago: Richard D. Irwin, Inc., 1945. Pp. 307.
8. Gray, W. S., and Leary, B. E. *What makes a book readable, with special reference to adults with limited reading ability: An initial study.* Chicago: University of Chicago Press, 1935. Pp. 358.

8a. Heron, A. R. *Sharing information with employees.* Stanford: Stanford University Press, 1942. Pp. 204.

9. Horn, E. *A basic writing vocabulary, 10,000 words most commonly used in writing.* Iowa City: University of Iowa, College of Education, Monographs in Education, First Series, No. 4. April, 1926.
10. Kitson, H. D. *The mind of the buyer.* New York: Macmillan Co., 1921. Pp. 211.
11. Maier, N. R. F. *Psychology in industry.* Boston: Houghton Mifflin Co., 1946. Pp. 463.
12. Mayo, E. *The social problems of an industrial civilization.* Andover, Mass.: The Andover Press, 1945. Pp. 150.
13. Oakley, C. A. *Men at work.* London: Hodder and Stoughton, Ltd. 1946. Pp., 301.
14. Ogburn, W. F. *Social change with respect to culture and original nature.* New York: The Viking Press, Inc., 1922. Pp. 365.
15. Pigors, P., and Myers, C. A. *Personnel administration. A point of view and a method.* New York: McGraw-Hill Book Co., 1947. Pp. 553.
16. Poffenberger, A. T. *Psychology in advertising.* Second Edition. New York: McGraw-Hill, 1932. Pp. 634.
17. Poffenberger, A. T. *Principles of applied psychology.* New York: D. Appleton-Century Co., 1942. Pp. 655.
18. Powell, L., and Schild, H. W. *How to prepare and publish an employee manual.* Third Edition. New York: American Management Association, 1947. Pp. 35.
19. Roethlisberger, F. J. *Management and morale.* Cambridge: Harvard University Press, 1944. Pp. 194.
20. Roethlisberger, F. J., and Dickson, W. J. *Management and the worker.* Cambridge: Harvard University Press, 1939. Pp. 615.
21. Scott, W. D., Clothier, R. C., Mathewson, S. B., and Spriegel, W. R. *Personnel management.* Third Edition. New York: McGraw-Hill Book Co., 1941. Pp. 589.
22. Smith, B. L., Lasswell, H. D., and Casey, R. D. *Propaganda, communication, and public opinion. A comprehensive reference guide.* Princeton: Princeton University Press, 1946. Pp. 435.
23. Thorndike, E. L. *A teacher's word book of the twenty thousand words found most frequently and widely in general reading for children and young people.* New York: Teachers College Bureau of Publications, Columbia University, 1931. Pp. 182.
24. Tiffin, J. *Industrial psychology.* Second Edition. New York: Prentice-Hall, Inc., 1947. Pp. 553.
25. Watkins, G. S., and Dodd, P. A. *The management of labor relations.* New York: McGraw-Hill Book Company, Inc., 1938. Pp. 780.
26. Yoder, D. *Personnel management and industrial relations.* Revised Edition. New York: Prentice-Hall, Inc., 1942. Pp. 848.

College Students' Opinions of Radio Advertising

Gove P. Laybourn and H. P. Longstaff

University of Minnesota

One important measure of the effectiveness of a particular kind of advertising is consumer acceptance of that advertising. Since radio advertising has grown to such importance, it was thought that the opinions of a group of students studying advertising might be informative. Thus the purpose of this study was to discover the likes and dislikes of a group of college students for various types of radio advertising.

The subjects were two hundred University of Minnesota Juniors and Seniors enrolled in the course, *Psychology of Advertising*. The majority of these students were advertising majors. Admittedly this was a highly selected group, nevertheless they represent not only consumers of their particular class but also the future leaders in the advertising profession and as such their opinions should be significant.

These students were asked to submit examples of radio advertising they liked best and the kind they disliked most. They were also instructed to give reasons supporting both their likes and dislikes. Some students submitted several advertisements which represented their likes and dislikes, but in order to maintain comparableness of the responses of all the subjects, only those advertisements which were given first choice were tabulated.

The data were tabulated so as to indicate the total number of responses favorable or unfavorable to each advertisement and these figures were converted into percentages. These data are presented in Tables 1 through 5. Tables 1 and 2 show the specific radio advertisements which were selected as the best liked and the most disliked, respectively. In Table 3 the general classes of radio advertisements which were selected as the best liked and the most disliked are given. Table 4 yields a comparison of the best liked and the most disliked commercials in that class of radio advertisements which was both best liked and most disliked. In Table 5 the reasons for liking and disliking radio advertisements as determined by this study and by a study made by the National Opinion Research Center are listed.

Referring to Table 1, it may be seen that the first four advertisements listed comprise the best liked of 42.5 per cent of the subjects. The others

Table 1

Best Liked Radio Advertisements Submitted by 200 College Students

Program Using Advertisement	Sponsor	Feature Liked	Number	Per Cent
Fibber McGee and Molly	Johnson Wax	Fibber's kidding of announcer	30	15.0
Jack Benny	Lucky Strike	Jack's frustration by quartet	26	13.0
Bob Hope	Pepsodent	"Poor Miriam" song	16	8.0
Station-break spot	United Fruit	"I'm Chiquita Banana"	13	6.5
Sunday Evening Hour	Ford	Informative institutional talk	8	4.0
Supper Club	Chesterfield	"ABC" theme song	7	3.5
Hollywood Players	Cresta Blanca	"C-r-e-s-t-a B-l-a-n-c-a" theme	6	3.0
Station-break spot	Virginia Dare	"Say It Again, Virginia Dare"	5	2.5
Station-break spot	Wildroot	"Wildroot Cream Oil Charlie"	5	2.5
Henry Morgan	Eversharp	Morgan's ribbing of sponsor	5	2.5
N. Y. Philharmonic	U. S. Rubber	Brief commercial at start and end	5	2.5
Hit Parade	Lucky Strike	Auctioneer and repeated slogans	5	2.5
Family Hour	Prudential	Informative talk on insurance	4	2.0
Bob Hawk	Camel	"C-a-m-e-l-s" and "Lemac" songs	3	1.5
Information Please	Parker	Brief and direct commercial	3	1.5
Telephone Hour	Bell	Interesting and informative talk	3	1.5
Cavalcade of Sports	Gillette	Commercial at sport-cast break	3	1.5
Joan Davis	Swan	Commercial in dialogue by cast	3	1.5
Blondie	Colgate	Commercial in dialogue by cast	2	1.0
Merideth Wilson	Canada Dry	"Canada Dry Water" theme song	2	1.0
Jack Carson	Campbell	"Mmm-mmm, Good" theme song	2	1.0
Others			34	17.0
No likes			10	5.0
			200	100.0

are liked by relatively much smaller numbers of subjects. Five per cent of the subjects state that they know of no particular radio advertisement which they like.

Referring to Table 2, it may be seen that the radio advertisement which is most disliked by the greatest number of subjects is the "Only Fifty Cents Down on a Lay By" skit, a station-break spot sponsored by the Northwestern Woolen Company.

It is interesting to note that this commercial, which was selected by more subjects than any other in this investigation and which provoked more unfavorable comment than all the others combined, failed to impress the name of its sponsor on almost one-third of the subjects who selected it. While three of these subjects admitted that they could not recall the name of the sponsor, the rest offered such variations as the following: Minnesota Woolen Mills, North Side Woolen Company, North Star Woolen Mills, North West Woolen Company, Northern State Mills, Northern Woolens, and Northwest Wool Company.

Table 2

Most Disliked Radio Advertisements Submitted by 200 College Students

Program Using Advertisement	Sponsor	Feature Disliked	Number	Per Cent
Station-break spot	N. W. Woolen	"Only 50¢ Down on a Lay By" skit	39	19.5
Hit Parade	Lucky Strike	"LS/MFT," auctioneer, and slogans	25	12.5
Jack Smith	Oxydol	"Oooh, That Oxydol Sparkle"	13	6.5
Station-break spot	Balm Barr	"I Use Balm Barr Lotion"	12	6.0
Station-break spot	Whiz	"Whiz, Best Nickel Candy"	10	5.0
Station-break spot	Lifebuoy	"B. O." foghorn effect	7	3.5
Red Skelton	Raleigh	"Proof Positive" slogans	7	3.5
Cavalcade of Sports	Gillette	Commercials interrupt sport-cast	6	3.0
Station-break spot	Super Suds	"Super Suds, Super Suds"	5	2.5
Station-break spot	Duz	"D-U-Z Does Everything"	5	2.5
Station-break spot	Virginia Dare	"Say It Again, Virginia Dare"	5	2.5
Station-break spot	Wildroot	"Wildroot Cream Oil Charlie"	5	2.5
Station-break spot	Ex-Lax	"Strike the Happy Medium"	4	2.0
Station-break spot	Pepsi Cola	"Pepsi Cola Hits the Spot"	3	1.5
Station-break spot	Grape Nuts	"Grape Nuts Flakes Are Good"	3	1.5
Station-break spot	Rinso	"Rinso White, Rinso White"	3	1.5
Cedric Adams	Taystee	Commercials interrupt news-cast	3	1.5
Radio Theater	Lux	Endorsements by guest stars	3	1.5
Station-break spot	Prell	"P-R-E-L-L, Prell Shampoo"	2	1.0
Bob Hawk	Camel	Medical claims	2	1.0
Screen Guild	Lady Esther	Lady Esther's whispering voice	2	1.0
Others			36	18.0
No dislikes			0	0.0
			200	100.0

The first five commercials comprise the most disliked radio advertisements of 49.5 per cent of the subjects. The others dwindle in unpopularity very rapidly.

It is rather interesting to observe that, while five per cent of the subjects are unable to mention a radio advertisement which they like, none of the subjects are unable to mention a radio advertisement which they dislike.

In Table 3, the best liked and most disliked radio advertisements are grouped together according to class or type of commercial. Examples of each class of advertisement are given under each main class or type.

At first glance, it may seem rather paradoxical that the type of commercial which is best liked by the greatest number of subjects should also be most disliked by the greatest number of subjects. This phenomenon may be understood by comparing the best liked and most disliked commercials of this type. Such a comparison is made in Table 4.

Table 3

Best Liked and Most Disliked Classes of Radio Advertisements

Best Liked Classes	No.	Per Cent	Most Disliked Classes	No.	Per Cent
Attention-Getting Commercials			*Attention-Getting Commercials*		
Presented as Jingles, Slogans, etc.	75	37.5	*Presented as Jingles, Slogans, etc.*	170	85.0
Pepsodent "Poor Miriam" song, United Fruit "Chiquita Banana," etc.			N. W. Woolen "Only 50¢ Down" skit; "LS/MFT," auctioneer, slogans, etc.		
Clever or Humorous Commercials			*Frequent or Lengthy Commercials*		
Incorporated into Dialogue by Cast	73	36.5	*Interfering with Rest of Program*	13	6.5
Johnson Wax Fibber McGee and Molly, Lucky Strike Jack Benny, etc.			Gillette Cavalcade of Sports, Taystee Bread News Time, etc.		
Simple and Direct Commercials			*Commercials Using Testimonials*		
Segregated from the Rest of Program	19	9.5	*or Medical Endorsements and Claims*	7	3.5
U. S. Rubber N. Y. Philharmonic, Parker Pen Information Please, etc.			Lux Radio Theater Camel Bob Hawk Show, etc.		
Serious Commercials Presented as			*Gushing Commercials Presented in*		
Informative or Interesting Talks	17	8.5	*Confidential or Emotional Tones*	6	3.0
Ford Sunday Evening Hour, Prudential Family Hour, etc.			Lady Esther Screen Guild Players, etc.		
Other Types of Commercials	6	3.0	*Other Types of Commercials*	4	2.0
No Likes	10	5.0	*No Dislikes*	0	0.0
	200	100.0		200	100.0

Table 4

Comparison of Best Liked and Most Disliked Commercials in that Class of Radio Advertisements which was Both Best Liked and Most Disliked

Attention-Getting Commercials Presented as Jingles, Slogans, etc.	Best Liked		Most Disliked	
	No.	Per Cent	No.	Per Cent
N. W. Woolen "Only 50¢ Down" skit	0	0.0	39	19.5
"LS/MFT," auctioneer, slogans	5	2.5	25	12.5
"Oooh, That Oxydol Sparkle"	1	0.5	13	6.5
"I Use Balm Barr Lotion"	0	0.0	12	6.0
"Whiz, Best Nickel Candy"	0	0.0	10	5.0
Lifebuoy "B. O." foghorn effect	0	0.0	7	3.5
Raleigh "Proof Positive" slogan	0	0.0	7	3.5
"Say It Again, Virginia Dare"	5	2.5	5	2.5
"Wildroot Cream Oil Charlie"	5	2.5	5	2.5
"Super Suds, Super Suds"	1	0.5	5	2.5
"D-U-Z Does Everything"	1	0.5	5	2.5
Ex-Lax "Strike the Happy Medium"	0	0.0	4	2.0
"Pepsi Cola Hits the Spot"	1	0.5	3	1.5
"Grape Nuts Flakes Are Good"	1	0.5	3	1.5
"Rinso White, Rinso White"	1	0.5	3	1.5
"P-R-E-L-L, Prell Shampoo"	0	0.0	2	1.0
United Fruit "Chiquita Banana"	13	6.5	1	0.5
Pepsodent "Poor Miriam" song	16	8.0	0	0.0
Chesterfield "ABC" theme	7	3.5	0	0.0
"C-R-E-S-T-A B-L-A-N-C-A"	6	3.0	0	0.0
"C-a-m-e-l-s" and "Lemac" songs	3	1.5	0	0.0
"Canada Dry Water" theme	2	1.0	0	0.0
Campbell "Mmm-mmm, Good" theme	2	1.0	0	0.0
Others	5	2.5	21	10.5
	75	37.5	170	85.0

An inspection of this table reveals that an attention-getting commercial which is selected by a certain number of subjects as being their best liked radio advertisement will not be selected by a comparable number of subjects as being their most disliked radio advertisement, and vice versa. For example, the Northwestern Woolen Company "Only Fifty Cents Down on a Lay By" skit was selected by 19.5 per cent of the subjects as their most disliked radio advertisement; not a single subject selected this commercial as his best liked radio advertisement. The commercials used on the Lucky Strike Hit Parade were selected by 12.5 per cent of the subjects as their most disliked radio advertisements; only 2.5 per cent—one-fifth as many—selected these commercials as their best liked radio advertisements.

Thus, a commercial which is best liked by many is most disliked by

few, and vice versa. In other words, the subjects are specific in their likes and dislikes: while some may be more favorably or less favorably disposed toward the general class of attention-getting commercials than others, most have some specific attention-getting commercials which they like and others which they dislike.

In Table 5 a comparison is made of the reasons for liking and disliking radio advertisements as determined by this study and by a study made by the *National Opinion Research Center*.[1]

Table 5

Percentage of Reasons for Liking and Disliking Radio Advertisements as Determined by this Study (N = 200) and by the N. O. R. C. Study (N = 2246)

Reasons for Liking and Disliking Commercials[1]	This Study		N. O. R. C. Study	
	Liked	Disliked	Liked	Disliked
Identifying slogans or sound effects	3%	42%	1%	3%
Singing or rhyming commercials	24	28	5	11
Variety vs. monotony	4	22	3	7
Clever humor vs. silly humor	31	17	8	4
Voice or manner of speaker or singer[2]	2	16	3	2
Fits program vs. interrupts program	34	13	10	2
Dignity vs. poor taste	8	10	2	4
Instructive vs. useless	6	9	4	*
Brevity vs. lengthiness	12	6	10	6
Unbiased vs. biased	2	6	2	4
Miscellaneous	1	3	5	4

[1] More than one reason per subject was possible.

[2] In N. O. R. C. study this category included only voice or manner of *announcer*.

* Less than half of one per cent.

The differences in relative and absolute magnitude of the percentages obtained in these two studies may be attributed to two factors.

The first factor, which possibly affected the differences in relative magnitude of the percentages to a greater extent than the absolute, concerns the composition of the groups of subjects involved. In this study the subjects were selected from an extremely narrow segment of the population: this group is markedly homogeneous as to age, intelligence, education, socio-economic status, and so forth. On the other hand, in the N. O. R. C. study the respondents were selected in such a way as to represent an accurate cross-section of the general population.

The second factor, which probably affected the differences in absolute magnitude of the percentages to a greater extent than the relative, con-

[1] Lazarsfeld, P. F., and Field, H. *The people look at radio.* Chapel Hill: The University of North Carolina Press, 1946.

cerns the collection of the data. In this study each subject was directed to state which radio advertisement he liked the best and which he disliked the most and to give the reasons for his choices. In the N. O. R. C. study each respondent was asked the following questions: 1. "Can you give me an example of what you think is the best advertising you've heard on the radio?" and 2. "Can you give me an example of what you think is the worst advertising you've heard on the radio?" If "Yes": "What didn't you like about it?" Of the 2246 respondents, only 43 per cent could give examples of what they liked and only 39 per cent could give examples of what they disliked. Therefore, the percentage of the total number of respondents giving reasons for their likes and dislikes is necessarily much smaller in the N. O. R. C. study than in this study.

Summary and Conclusions

1. The best liked classes of radio advertisements are the attention-getting commercial presented as a jingle, slogan, skit, and so forth and the clever or humorous commercial incorporated into the dialogue by the cast.

2. The most disliked class of radio advertisements is the attention-getting commercial presented as a jingle, slogan, skit, and so forth.

3. An attention-getting commercial which is selected by a certain number of subjects as their best liked radio advertisement will not be selected by a comparable number of subjects as their most disliked radio advertisement, and vise versa. While many attention-getting commercials are generally disliked, certain others are generally liked.

4. The principal reasons given for liking radio advertisements are the following: (1) fits program, (2) clever humor, (3) singing or rhyming commercials, and (4) brevity.

5. The principal reasons given for disliking radio advertisements are the following: (1) identifying slogans or sound effects, (2) singing or rhyming commercials, (3) monotony, (4) silly humor, and (5) voice or manner of speaker or singer.

6. In view of the above it seems that great care should be exercised in composing radio advertising and that it should then be carefully pretested. It would be difficult indeed to justify some of the more highly disliked radio advertisements.

Received April 1, 1947.

Opinion Polling with Mark-Sensed Punch Cards

N. L. Gage and H. H. Remmers

Purdue University

Ever since psychologists began studying individual differences, they have been plagued with the difficulties of handling large masses of data. The greater the number of cases studied, and the more intensive the kinds of statistical analyses attempted, the greater have been the burdens of clerical labor to be carried in the search for valid generalizations.

One early step in the direction of easing this burden was the development of objectively scorable, or short-answer, tests whose scoring could be turned over to clerks. But with the development of large scale, periodic testing programs, it soon became evident that further aid was desirable to reduce the work in scoring thousands of test booklets. To meet this need, many kinds of scoring devices were developed of which the most successful and widely used has probably been the International Business Machines Test Scoring Machine (IBM). With these devices, the rapid and accurate scoring of large numbers of tests became much more economical and efficient. By-products of these devices were the more extensive utilization of item analysis techniques and also, of more dubious value, a greater emphasis on the multiple-choice type of test item.

For one kind of appraisal of psychological data, however, the test scoring machines have distinct limitations. This is the situation, of which opinion polling may be the prime example, where we are interested not in securing a single total score based on a large number of responses, but rather in counting the frequency of occurrence of single responses and of tabulating these frequencies in relationship both to other single questions and to various types of personal data. This is the kind of analysis that is most frequently made of the data secured by opinion polling and market research surveys.

The trends in opinion polling toward depth interviewing, the clinical approach, open-end questioning, and other more intensive and subjective techniques, may become accentuated. But it is probable that for some time to come opinion polling will continue to depend largely on questions the responses to which are restricted to a small number of alternatives. Furthermore, as opinion pollers become more interested in making the

higher order breakdowns that are necessary for unequivocal investigations of the relative importance of determinants of opinion, they will need samples consisting of much larger numbers of cases than they have hitherto required. Thirdly, in some kinds of polling situations, such as that in which the *Purdue Opinion Poll for Young People* finds itself, it is necessary to secure samples in the form of intact groups, say whole classrooms or schools, which require larger numbers of cases for a given level of reliability than do samples that are not composed of intact groups. All these considerations point to the desirability of further aids in the analysis of data. The purpose of this paper is to describe the use for such purposes of the mark-sensing attachment to the IBM Reproducing Punch.

What the mark-sensing attachment does is to convert pencil marks on an IBM card to punched holes in the card at a speed of 100 cards per minute. During the past year, the *Purdue Opinion Poll for Young People* has conducted three polls by means of the mark-sensing technique. The following details of the procedure are pertinent here.

1. When used for mark-sensing, the IBM card is divided into twenty-seven columns, each covering three of the columns in which holes can be punched. The pencil mark must be made with an electrographic pencil so as to cover at least two of the punching columns. An electrical impulse conducted by the pencil mark is then passed through an electronic amplifying unit to activate the mechanism that punches the hole in the card. Ten mark-sensing brushes are available so that ten mark-sensing columns can be punched with one passage of the card through the machine. By means of a flexible wiring arrangement the hole can be punched from any mark-sensing column to any other column on the card including the column in which the mark is placed. Usually, the machine is wired so as to punch the hole in the same column in which the mark has been made.

2. In order to make the same card form usable in many different polls or rating schemes, the card was designed so that each of the 324 marking spaces was uniquely numbered. The various response alternatives furnished with the questions of opinion and personal data were assigned numbers to correspond with certain spaces on the card. The pupil is instructed to read the question and choose the answer that applies to himself, notice the number of that answer, find the space with that number on the card and fill in that space with a heavy mark of the special pencil.

3. The mark-sensing device will punch any or all of the twelve spaces in a given column with one passage of the card through the machine. It is consequently possible to exploit the possibilities of multiple-punching to a much greater extent than is feasible with manual key-punching.

That is, although in manual punching it is inconvenient to punch more than a single item in any given column, it is as easy with mark-sense punching to secure twelve punches in a column as it is to secure a single punch. Since each of the questions on the Poll had three response alternatives, it was possible to design the coding so that four three-choice questions were coded for a single column. Such multiple punching had distinct advantages when it came to counting the responses to the questions. By means of the IBM counting sorter, the responses to four three-choice questions could be counted simultaneously with one passage of the cards through the sorter. The number of pupils choosing each of the three answers to the four questions was then recorded by hand from the Veeder counters on the sorter. Since the sorter handles 400 cards per minute and yields twelve frequencies per run, the use of multiple-punching was four times more efficient than if a single-punching code had been used.

4. Multiple-punching also had advantages in the process of securing scatter diagrams for inter-item correlation. The sorter was used to divide the pupils into three groups according to the three responses to a given question. These three groups were then sent through the sorting machine separately and the numbers in each group who had chosen each of the three responses to the four questions in a given column were counted simultaneously. The multiple-punching in effect reduced the number of runs required for securing intercorrelations to one-fourth the number otherwise necessary.

5. In performing a scale analysis by the Cornell technique as described by Guttman,[1] the IBM tabulating machine equipped with digit selectors also proved much more efficient by virtue of the multiple-punching code. Preliminary weights were assigned to each of the response alternatives. Then the cards were assigned total scores by using each card as a window stencil placed over another IBM card marked as a scoring key. The total scores were then punched in the card, and the cards were sorted to place them in descending order of total scores. When the cards were sent through the tabulator, the responses to four items at a time were listed. Inspection for unidimensionality and reproducibility could then be carried out in accordance with the Cornell technique.[2]

Because of the novelty of the method, it was important to ascertain the amount of error resulting from pupils' failures to comprehend and follow the directions for recording responses on the cards. In the situations under which these polls were administered, by teachers located in widely scattered schools and instructed only by means of a sheet of directions,

[1] Guttman, L. The Cornell technique for scale and intensity analysis. *Educational and Psychological Measurement*, 1947, 7, 247–279.

[2] Gage, N. L. *Scaling and factorial design in opinion poll analysis.* Lafayette, Ind.: Purdue University, Studies in Attitudes, X, 1947, 84 pp.

it was found that between one and two per cent of the marks had been made improperly. It is anticipated that with improved directions to pupils and with increasing familiarity of pupils with this method, this error can be further reduced.

Further applications of the mark-sensing technique can readily be made. In general, wherever interest is centered not on a total score based on a number of questions but rather on the frequency of various kinds of responses, the mark-sensing technique may have great advantages in the analysis of large numbers of cases. Among the other applications that have already been considered are the summarization of ratings of a single person by a large number of other persons, as when classroom teachers are rated by their pupils. Polls carried out by interviewers can also be analyzed in this way if the interviewers record the responses of interviewees on mark-sensing cards rather than on tally sheets. The scaling of items by the Thurstone equal-appearing intervals technique can be done by having the judges mark their ratings of items on cards rather than by having them place the items in various piles.

Summary

In summary, then, we have described the application of mark-sensed punch card technique to the recording and analysis of large masses of psychological data. Application of the technique to questionnaire studies as in the *Purdue Opinion Poll for Young People* has proved successful in terms of speed and accuracy of tabulation, and in terms of increased efficiency in securing item intercorrelations and performing scale analyses. In general, the method will bear investigation of its usefulness in a wide variety of psychological investigations.

Received May 8, 1947.

News Notes

The name "Committee on Aviation Psychology" has been substituted for "Comittee on Selection and Training of Aircraft Pilots" to designate a committee of the Division of Anthropology and Psychology, National Research Council, which has conducted research in the field of aviation psychology since 1939. As in the past, the work of the Committee is supported with funds allotted by the Civil Aeronautics Administration, although steps have recently been initiated to undertake research for other Government agencies.

The Committee has from its very beginning conducted research involving the maintenance as well as the selection and training of aircraft pilots. Among the more than 70 reports published by the Committee in the Technical Series, Division of Research, Civil Aeronautics Administration, there are many which are concerned with the psychological aspects of fatigue, accidents, air sickness, etc.

The current research program of the Committee on Aviation Psychology is largely concerned with air transport pilots and, most particularly, with human factors in airplane accidents. The program includes research on the selection, upgrading, and certification of pilots; studies of stall warning devices; psychological aspects of instrumentation; as well as investigation on methods of training civilian and commercial pilots. Such research is being conducted through The Ohio State University, the American Institute for Research, the Educational Research Corporation, and other agencies. The Executive Subcommittee will be pleased to consider proposals for grants-in-aid to research personnel working in universities and other institutions who are interested in carrying out investigations in the field of aviation psychology under the auspices of the Committee. Proposals should be submitted to Dr. M. S. Viteles, Chairman, National Research Council Committee on Aviation Psychology, University of Pennsylvania, Philadelphia, Penna.

The present membership of the Committee on Aviation Psychology includes:

Brigadier General Milton W. Arnold, Vice President—Operations and Engineering, Air Transport Assn. of America; *Comdr. Norman L. Barr, Division of Aviation Medicine, Bureau of Medicine and Surgery, U. S. Navy; *Dr. George K. Bennett (ex officio), President, Psychological Corporation; *Dr. Dean R. Brimhall, Assistant to the Administrator for Research, Civil Aeronautics Administration; *Dr. Paul M. Fitts, Chief, Psychology Branch, Aero Medical Laboratory, Wright Field; *Dr. Frank A. Geldard, Professor of

Psychology, University of Virginia; Captain B. Groesbeck, Division of Aviation Medicine, Bureau of Medicine and Surgery, Navy Department; Major General Malcolm C. Grow, Air Surgeon, Army Air Forces; George E. Haddaway, Editor, Southern Flight, Air Review Publishing Corporation; *Dr. A. I. Hallowell (ex officio), Chairman, Division of Anthropology and Psychology; Professor of Anthropology, University of Pennsylvania; Dr. J. G. Jenkins, Professor of Psychology; Head, Department of Psychology, University of Maryland; *Captain Wilbur E. Kellum, U. S. Navy, School of Aviation Medicine, Pensacola, Florida; Dr. Peter C. Kronfeld, Associate Professor of Ophthalmology and Director of Education, Illinois Eye and Ear Infirmary, University of Illinois; Jerome Lederer, Chief Engineer, Aero Insurance Underwriters; *Dr. Donald B. Lindsley, Professor of Psychology, Northwestern University; Dr. W. R. Miles, Professor of Psychology, School of Medicine, Yale University; Dr. C. L. Shartle, Professor of Psychology, The Ohio State University; *Lt. Col. Anthony C. Tucker, Office of the Air Surgeon, Headquarters, Army Air Forces; and *Dr. M. S. Viteles, Chairman, Committee on Aviation Psychology; Professor of Psychology, University of Pennsylvania.

Personnel Psychology, a Journal of Practical Research, has announced its first issue for January, 1948. It will publish papers related to research in the personnel field, particularly in selection, placement and evaluation, training, job analysis, classification and compensation, employee relations, motivation and morale, industrial design, and work conditions. Readability is to be combined with scientific excellence as a criterion.

Personnel Psychology will be edited by G. Frederick Kuder, Duke University, with M. W. Richardson, Richardson, Bellows, Henry and Co.; George K. Bennett, Psychological Corporation; and W. V. D. Bingham as advisors on editorial policy.

Manuscripts and editorial communications should be sent to Dr. Frederick Kuder. If sent before February 1, 1948, they should be addressed to 306 Edmonds Building, 915 Fifteenth Street, Washington 5, D. C. After February 1, 1948, Dr. Kuder's address will be Department of Psychology, Duke University, Durham, North Carolina.

Personnel Psychology will be issued quarterly, beginning January, 1948. The yearly volume will consist of 400–500 pages. The subscription rate is $6.00 per volume, Foreign, $7.00 per volume, single copies $2.00. Subscriptions and business communications should be sent to Dr. Erwin K. Taylor, 1727 Harvard Street, N. W., Washington 9, D. C.

* Member of Executive Subcommittee.

Book Reviews

Maier, Norman R. F. *Psychology in industry: a psychological approach to industrial problems.* Boston: Houghton Mifflin, 1946. Pp. xvi + 463. $3.00.

This kind of book has been long overdue. It is a systematic but non-technical presentation of the applications of psychological principles and techniques to the functioning of industrial organizations. As such, it is "primarily intended for the many who are concerned with human problems in industry and who themselves are not industrial psychologists." It should find wide use in college courses aimed at acquainting industry's potential executives and technicians with psychology's significance for their future vocational activities and responsibilities.

Dr. Maier places strong emphasis throughout his book on the dynamics of human behavior in industry, specifically on matters of motivation, attitudes, and morale. Five of the nineteen chapters are devoted exclusively to problems of attitudes, frustration, morale, and motivation. The systematic viewpoint taken is that "frustration and motivation represent two distinct sources of action in the individual" and that "to understand the behavior at any particular time we must know which mechanism is dominating." The chapter on "Frustration as a Factor in Forming Attitudes and Developing Social Movements" is of especial significance. However, that significance might be even clearer if the detailed discussion of motivation preceded the discussion of frustration instead of following several chapters later.

Two introductory chapters discuss the scientific study of behavior and the concept of causation in behavior. Other chapters cover major topics in the field of industrial psychology in a thorough but non-technical way. These topics include the measurement of proficiency, the nature and use of tests, motion-and-time study, training and learning, fatigue, accidents, working conditions, and labor turnover. The basic frame of reference—attitudes, morale, and motivation—is considered in relation to each of these topics. Thus each technique and principle is seen against the background of a living, functioning industrial organization. The concluding chapter reviews and points up the significance of each major topic for, first, the first-line supervisor; then, the employee counselor; and finally, higher levels of management. Following the last chapter is a chapter by chapter bibliography.

The book contains a considerable amount of material that is relatively new to the field of industrial psychology. This material includes discussions of the level of aspiration, of resistance to interruption of a task, and of the application of sociometry to industrial situations, and preliminary reports of experiments conducted by Alex Bavelas on democratic techniques of supervision in industry.

There are a few serious omissions. The concept and problem of reliability are not considered at all in connection with either merit rating or tests. This concept is no more technical than others included in the book, and certainly it is of major importance. Although there is an excellent treatment of training and learning in industry, there is no mention of the *Training Within Industry* program. In spite of the great emphasis on morale, the morale of first-line supervisors, which is currently a very live issue, is barely mentioned. The suggestion in the final chapter that higher management doesn't need as high a level of supervisory skills as first-line supervisors would be difficult to defend.

This reviewer's knowledge of and experience with first-line supervisors in several kinds of business and industrial organizations convinces him that the book is much better adapted to college students and college-trained technicians and supervisors than to the average first-line supervisor in industry. It is pitched at too high a level of comprehension for a large proportion of industry's present first-line supervisors. And it is written too much in the style of the typical college text to have strong appeal for such supervisors.

Several experiments with animals are cited as illustrations of various mechanisms. To expect the average personnel manager, industrial engineer, or supervisor to believe that such experiments have significance for his work seems to be a bit naive. Practically no efforts are made in the book to suggest why animal experiments are relevant.

For an introductory college course on psychology's applications for business and industry *Psychology in industry* is an excellent text. A non-technical book on industrial psychology that will capture the interest of industry's first-line supervisors without sacrificing accuracy is still to be written.

Kenneth A. Millard

Macalester College,
St. Paul, Minnesota

Gray, J. Stanley, et.al. *Psychology in human affairs.* New York, McGraw-Hill, 1946. Pp. viii + 646. $3.75.

The reader may recall a similar book edited by the same author in 1941 with contributions by a dozen psychologists. The present work is not a revision and, in fact, the earlier one is not mentioned. However,

they do cover much the same ground. Gray does several chapters himself and with one exception the collaborators are different from those who participated in the previous enterprise. Most of the present collaborators are less well known but, on the whole, the work is better integrated. Perhaps the editor was in a better position to direct and modify their contributions and perhaps there was some advantage in the fact that many of them were geographically accessible. At any rate, there is less apparent discontinuity than in the average book by a conglomeration of authors. There are still a few important omissions (infra) presumably where one contributor expected another to cover the point, and there are apparent differences in style, but neither of these shortcomings is too serious. After reading the first six chapters the reviewer wished that Gray had kept right on and done all of them.

The rationale back of the arrangement of the chapters is not entirely clear. While business and industrial problems are discussed in sequence, the chapters on adjustment, mental illness, and clinical practice are not contiguous. However, most of the material is there anyway, which is the important thing. The major fields of education, business, and medicine in their psychological implications are treated adequately, with the field of criminal and legal psychology slighted somewhat. There are also chapters on public opinion, music, art and literature, and military psychology.

The book is quite definitely a textbook, presumably aimed at about the second course after the student has a general background in psychology. The treatment is commendably factual, citing experiments, and with extensive annotation. This would make it a bit heavy for the layman but is altogether appropriate in a text.

Any reviewer is more favorably impressed by some parts of a book than by others. In the present case, vocational guidance seems to be Gray's best chapter and the one on public opinion is among the best from the contributing authors. The discussion of psychology in crime appears to be the least satisfactory and contains rather little psychology.

There are a few serious omissions. There is nothing on scientific crime detection, a field in which the public is rather misinformed and in which students might profitably be set straight. There is nothing about the psychology of testimony. Another major omission is morale in industry and the role of psychological factors in supervision. Morale is mentioned in the military discussion but not in the industrial. In employment psychology, there is all too brief a presentation of employment tests, with a disproportionately extensive treatment of interviews and personal data. Tests constitute psychology's greatest contribution in this field. There are some other places where there is not an actual omission but where the discussion might profitably be extended a little.

Points which impressed the reviewer in this respect are how to study, personality tests, psychotherapy, and clinical techniques.

On the other hand, there are some points where the treatment is more extensive than is usually the case in texts on applied psychology. Speech correction gets a whole chapter, and the psychological effects of drugs has a fairly detailed treatment. Neither of these detracts at all from the value of the work, but if it were a question of condensing something in order to include some of the omissions mentioned above, these topics might be considered.

The question of statistics is always a moot one in a book on applied psychology. It is not quite clear whether it is assumed that the student will have had some of the fundamentals before tackling this particular book. It practically assumes that he knows something about correlation coefficients because there is only a brief footnote regarding them when they first appear in the discussion. It is well for the student to get the notion that psychology is a quantitative science. The text will give him that much, but would be better if, in addition, he learned a little more about interpreting quantitative data in terms of variability, correlations, and probability concepts.

The book covers the ground about as well as any text could be expected to do in view of the terrific size of the field of applied psychology at the present time. It presents a very considerable amount of source material, citing facts in support of a great many of the points brought out and implementing the discussion with numerous tables and graphs. The annotations and footnotes would be helpful to anyone interested in running down detailed material. While the work is not actually a source book, it borders on it in spots. From the teaching standpoint, the outlines at the beginning of each chapter will be helpful, and the suggested readings at the ends of the chapters are well chosen. If a student gets through this book under a teacher who can find his way around in the field, he will certainly know something about applied psychology.

Harold E. Burtt

Ohio State University

Robinson, Francis P. *Effective study.* New York: Harper and Brothers, 1946. Pp. ix + 262. $3.00.

This extensive revision of the author's previous (1941) *Diagnostic and remedial techniques for effective study* is a manual or workbook written to be used by students enrolled in how-to-study courses. It could be used with equal profit in many so-called college orientation courses. The book represents a coordination of the diagnostic and training devices and principles which have been demonstrated to improve scholarship at

Ohio State University. It is designed to "free the instructor so that he may become a counselor rather than a task master."

The first part of the book describes "Higher Level Work Skills"; these are traditional how-to-study topics and principles treated in non-traditional fashion. The motivated student to whom the logic of its presentation appeals should benefit greatly from using the "tricks of the trade" which receive detailed treatment in this section. Remedial reading and effective language usage are described, and various exercises provided, in the second part. In the last part problem areas familiar to the student personnel worker receive considerably more attention than is customary in books designed to teach study skills.

Throughout the book are included charts, graphs, rating scales, check lists of various sorts, and other devices by which original status and progress throughout a course may be known to the student. The motivational value of knowledge of progress is wisely exploited by the author.

Objective tests cover fields ranging from knowledge of social usage or etiquette, through dictionary and library usage, to principles of mathematical formulation. Norms for (presumably Ohio State) freshmen and keys are included in the book. The diagnostic tests provide the student in a how-to-study course with opportunities to evaluate his skills and adjustments to college. It is on the basis of data made available through the use of the book that the student, under the guidance of the instructor, develops study skills and adapts them to courses currently being studied.

The author's appeals are frankly made to serious students of all levels of ability who appreciate the logic of doing a job right. This is not a book to be given as a "going away to college" present to a freshman. Its effective use demands guidance by a competent instructor, who, in addition, realizes it does not provide short cut methods for so-called group counseling. If used in a college having various agencies for remedial work and counseling to which the instructor could refer some of his students, the book will prove all the more valuable.

Use of the book for the purposes for which it was written commits the instructor selecting it as a work manual to a plan of action which may, indeed, reorient the instructor and school as well as the students.

C. d'A. Gerken

The State University of Iowa

Blankenship, Albert B. (ed.). *How to conduct consumer and opinion research. The sampling survey in operation.* New York: Harper & Brothers, 1946. Pp. xi + 314. $4.00.

Dr. Blankenship has induced twenty-five busy research practitioners (nine of them psychologists) to describe important examples of their work.

The resulting chapters constitute the best available collection of material on applications of questionnaire and interview research to business uses. The contributions focus predominantly on market research; only a small part of the volume deals with public opinion polling and with attitude research for government.

Emphasis is placed on the practical importance of the research described. The aim, the editor states, is to help businessmen determine what services they should utilize in this field and to acquaint students with research applications of the "sampling survey" which lie beyond "elementary knowledge of techniques."

The twenty-two chapters apart from the introduction may be grouped as follows: eight are on market research, including studies of product improvement, brand acceptance, dealer performance, and advertising; five are on radio research, principally audience measurement for commercial purposes; two deal with magazine reading; two with employee attitudes and public relations; four are on governmental use of sampling surveys and one is on public opinion polls.

A few of the more general chapters cover the purposes of marketing studies and the steps in questionnaire research, methods used in testing advertisements and in developing advertising copy, public opinion polling activities and the newer procedures in sampling. Other chapters sketch more specific examples of applied opinion work. The reader learns how radio audience size is measured in the famous Hooperatings; he is told the methods for counting magazine readers in *Life's* elaborate "Continuing Study of Magazine Audiences"; he sees the Psychological Corporation's "brand barometer" and the Industrial Surveys Company's use of extensive consumer panels to determine people's changing brand buying behavior and its relation to their magazine reading and radio listening; he looks at practical illustrations of studies to improve products and advertisements; he is introduced to the techniques of employee attitude measurement, research on magazine editorial problems, and detailed analysis of radio programs. Government survey work is represented by brief reports of studies by the Office of War Information, the Office of Civilian Requirements, the Bureau of the Census, and the Division of Program Surveys in the Department of Agriculture (a total of 44 pages for all of these).

The treatment throughout is severely practical and empirical. There is almost no reference to guiding theories or use of hypotheses—no suggestion that sound theory may also contribute valuably to applied research. There is likewise little critical discussion of techniques or comparative evaluation of methods. Problems of analysis and interpretation of findings are especially slighted. And research purposes are dealt

with somewhat idealistically, as if the selection of problems and procedures were always aimed at true problem solving and never at proving a case or securing sales ammunition. In short, the book is less a rounded and balanced treatise on research methods than it is an account of successful research services.

As such, it will be extremely useful in providing students of marketing and opinion research with an up-to-date view (if not quite complete and realistic) of the practical work being done. There has been remarkably little writing of this type available. Consequently the book fills a genuine need. Those who are working and teaching in this field owe a definite vote of thanks to the editor and contributors.

Arthur W. Kornhauser

Wayne University

Maslow, Paul. *The analysis and control of human experiences.* The individual seen through the Rorschach. Vol. I and II. Brooklyn, New York: Paul Maslow, 1946 and 1947. Pp. 195 and 229. $3.50 each.

The study of "Rorschach Psychology" differs from other scientific study, according to the author, in that ". . . the study of the personality structure is different because it cannot be seen, felt, smelled, tasted or touched and actually does not exist. 'It' is invisible, nonmaterial and without extension and weight. . . . The Rorschach operates by materializing the non-existent and measuring this 'substance' to transform the complete personality structure into an absolute and, in some instances, grim reality" (Vol. I, p. 2).

Accepting the Rorschach in the unqualified fashion illustrated above the first volume deals mainly with a series of short essays (50 chapters) in which a great variety of personality types and terms are analyzed in the language of the Rorschach. The author places a great deal of reliance upon the specific meaning of single determinants.

The second volume (58 chapters) consists largely of a series of essays on science, epistemology, religion, politics, ethics, therapy and a variety of miscellaneous topics. These are illustrated by means of Rorschach terms or clarified by knowledge of personality types which have been analyzed through the Rorschach.

The following quotation describing part of the author's technique of Rorschach analysis illustrates the style of the book ". . . Note the rhythm of the record, the arrangement of details, the attraction-repulsion patterns, the quality of the stimulation and the varying pitches of intensity. Watch for holes, dead-ends, contrasts, . . . brilliant bursts, persistent triviality, massive complexity, depth, width, laxness, unity. . . ." (Vol. II, p. 14). No referents for these descriptive phrases are provided.

The two volumes have no systematic organization. No data are presented and it is not clear to what extent the series of speculations are based upon the clinical experience or reading of the author. To those who accept the Rorschach as uncritically as the author does these volumes may contain some interesting speculations, to others they will not be profitable reading. The two volumes are mimeographed and paper bound. The bibliography contains 212 items.

Julian B. Rotter

Ohio State University

New Books, Monographs, and Pamphlets

Books, monographs, and pamphlets for listing and possible review should be sent to Donald G. Paterson, Editor, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota

Measurement of consumer interest. C. West Churchman, Russell L. Ackoff, and Murray Wax, Editors. Philadelphia: University of Pennsylvania Press, 1947. Pp. 214. $3.50.

Dvorine color discrimination screening test. Israel Dvorine. Baltimore: Israel Dvorine, 1947. Pp. 10. $2.50.

The incidence of neurosis among factory workers. Russell Fraser, et al. Report No. 90. Industrial Health Research Board, Medical Research Council. London, England: H. M. Stationery Office, 1947. Pp. 66. 1s. 3d.

Study your way through school. C. d'A. Gerken. Chicago: Science Research Associates, 1947. Pp. 48.

Psychotherapy in child guidance. Gordon Hamilton. New York: Columbia University Press, 1947. Pp. 340. $4.00.

Sociometry of leadership. Helen J. Jennings. New York: Beacon House, 1947. Pp. 28. $1.50.

Rehabilitation of the physically handicapped. Henry H. Kessler. New York: Columbia University Press, 1947. Pp. 274. $3.50.

The relationship between content of an adult intelligence test and intelligence test score as a function of age. Rose E. Kushner. New York: Bureau of Publications, Teachers College, Columbia University, 1947. Pp. 59. $2.10.

Mits, wits, and logic. Lillian R. Lieber. New York: W. W. Norton and Co., Inc., 1947. Pp. 240. $3.00.

The story of hypnotism. Robert W. Marks. New York: Prentice-Hall, Inc., 1947. Pp. 246. $3.00.

Psychological testing. James L. Mursell. New York: Longmans, Green and Co., Inc., 1947. Pp. 449. $4.00.

Readings in social psychology. Theodore M. Newcomb, et al. New York: Henry Holt and Co., 1947. Pp. 672. $3.85.

Abnormal psychology. James D. Page. New York: McGraw-Hill Book Co., Inc., 1947. Pp. 441. $4.00.

Personnel administration. Paul Pigors and Charles A. Myers. New York: McGraw-Hill Book Co., Inc., 1947. Pp. 553. $4.50.

Foundations for American education. Harold Rugg. Yonkers-on-Hudson, New York: World Book Company, 1947. Pp. 826.

The selected writings of Benjamin Rush. Dagobert D. Runes, Editor. New York: The Philosophical Library, 1947. Pp. 433. $5.00.

Introduction to methods in experimental psychology. Second Edition. Miles A. Tinker, New York: D. Appleton Century, Inc., 1947. Pp. 232. $4.00.

The thematic apperception test. Silvan S. Tomkins. New York: Grune and Stratton, 1947. Pp. 297. $5.00.

The reduction of intergroup tensions: a survey of research on problems of ethnic, racial, and religious group relations. Robin M. Williams. New York: Social Science Research Council, 1947. Pp. 153. $1.75.

So you want to help people. Rudolph M. Wittenberg. New York: Association Press, 1947. Pp. 174. $3.00.

Preparing an employee handbook. London, England: Institute of Personnel Management, 1947. Pp. 36. 2s. 6d.

Ready in March

AN INTRODUCTION TO CLINICAL PSYCHOLOGY

Edited by **L. A. PENNINGTON**
Associate Professor of Psychology, University of Illinois
and **IRWIN A. BERG**
Associate Professor of Psychology, Pomona College.

Prepared by 23 Contributing Specialists

THIS new book aims to present eclectically what clinical psychologists actually do. No single clinician deals regularly with all varieties of problems nor with all clinical methods considered, yet every clinical psychologist is professionally concerned with some of them. The scope of this volume is represented in varying proportion in the daily, collective activities of clinical psychologists.

Throughout, the editors have stressed the similarity of problems encountered in the field of clinical behavior, regardless of the occupational placement of the practicing psychologist. In the editors' opinion the problems remain much the same, although the purposes for which clinical procedures are applied may vary. The sole aim of this volume has been to present an introductory description of the problems encountered, and of the diagnostic and therapeutic techniques used by the qualified clinician at work.

As an introductory text it will provide broad familiarity with clinical problems and methods and will serve as a convenient springboard for subsequent, more specialized courses.

600 pages, **$5.00**

THE RONALD PRESS COMPANY
15 East 26th Street **New York 10**

JOURNAL OF APPLIED PSYCHOLOGY

YEAR	VOLUME	AVAILABLE NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
		MAR	JUN	SEP	DEC				
1917	1	–	–	3	4			$1.75	$3.50
1918	2	1	2	3	4			$1.75	$6.00
1919	3	1	2	3	4			$1.75	$6.00
1920	4	1	2 & 3		4			$1.75	$6.00
1921	5	1	2	3	–			$1.75	$5.25
1922	6	1	2	–	4			$1.75	$5.25
1923	7	–	2	3	4			$1.75	$5.25
1924	8	1	2	3	4			$1.75	$6.00
1925	9	1	2	3	4			$1.75	$6.00
1926	10	–	2	–	4			$1.75	$3.50
		FPB	APR	JUN	AUG	OCT	DEC		
1927	11	1	2	3	4	5	6	$1.25	$6.00
1928	12	1	2	3	4	5	6	$1.25	$6.00
1929	13	1	2	3	4	5	6	$1.25	$6.00
1930	14	1	2	3	4	5	6	$1.25	$6.00
1931	15	1	2	3	4	5	6	$1.25	$6.00
1932	16	1	2	3	4	5	6	$1.25	$6.00
1933	17	1	2	3	4	5	6	$1.25	$6.00
1934	18	1	2	3	4	5	6	$1.25	$6.00
1935	19	1	2	3	4	5	6	$1.25	$6.00
1936	20	1	2	3	4	–	6	$1.25	$6.00
1937	21	1	2	3	4	5	6	$1.25	$6.00
1938	22	1	2	3	4	5	6	$1.25	$6.00
1939	23	1	–	3	4	5	6	$1.25	$6.00
1940	24	–	–	3	4	5	6	$1.25	$5.00
1941	25	1	2	3	4	5	6	$1.25	$6.00
1942	26	1	2	3	4	5	6	$1.25	$6.00
1943	27	1	2	3	4	5	6	$1.25	$6.00
1944	28	1	2	3	4	5	6	$1.25	$6.00
1945	29	1	2	3	4	5	6	$1.25	$6.00
1946	30	1	2	3	4	5	6	$1.25	$6.00
1947	31	–	2	3	4	5	6	$1.25	$6.00
1948	32	By subscription, $6.00						$1.25	—
		List price, Volumes 1 through 31							$177.75
		30% Discount							53.33
		Net price, Volumes 1 through 31							$124.42

Information about the Journal of Applied Psychology: only 11 numbers are out of print in the thirty volumes. These are scattered throughout the years.

Volumes 1 through 10 had four numbers per year; after that there are six numbers per year. Number 3 of Volume 24 will be sold only with an order for several volumes.

Information about prices: the journal is uniformly priced at $6.00 per volume. When four numbers per volume were issued, the price per number is $1.75; when six numbers per year were issued, the price per number is $1.25. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over

20% on orders of $100.00 and over

30% on orders of $150.00 and over

Current subscriptions and orders for back numbers should be addressed to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 Massachusetts Avenue, N. W.

Washington 5, D. C.

Encyclopedia of
CHILD GUIDANCE

Edited by* RALPH B. WINN, *Ph.D.

"ALL the apparatus, labels and paraphernalia of the child psychologists have been assembled, described and defined in the *Encyclopedia of Child Guidance*. This volume of 456 pages, edited by Ralph B. Winn, brings together in alphabetical arrangement the efforts of more than sixty experts to analyze and characterize the technical terms employed in diagnosing children's mental traits.

"An excellent job has been done in thus bringing between two covers so large and varied an amount of information about one of the youngest and most important areas of educational advance."—*The Journal of Education.*

"THIS encyclopedia is the result of the combined work of more than 60 psychiatrists, educators and social workers. It is a volume of practical value to every person or group concerned with the training and development of children. The book deals with many phases of child guidance, psychology, education, social and clinical work. Because of its simple and clear presentation, it can be used by intelligent parents as well as professional workers.—*Elementary Education Bulletin.*

From the Table of Contents

★ Ability ★ Antisocial Behavior ★ Autistic Thinking ★ Behavior Problems ★ Case History ★ Characterology ★ Chronaxy ★ Conditioning ★ Crippled Children ★ Cumulative Records ★ Divorce Effects ★ Domination ★ The Ego ★ Examination, Psychiatric ★ Gifted Child ★ Instinct ★ Introversion ★ Lying ★ Mental Hygiene ★ Negativism ★ Parent-Teacher Relations ★ Play ★ Problem Tendencies ★ Racial Differences ★ Refugee Children ★ Religious Beliefs ★ Remedial Work ★ Rorschach ★ Sadism ★ Sex Education ★ Speech Disorders ★ Suggestion ★ Transference ★ War Effects

LIMITED EDITION **ORDER NOW** **$7.50**

PHILOSOPHICAL LIBRARY

Publishers

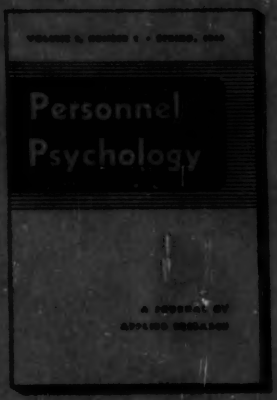
PHILOSOPHICAL LIBRARY, *Publishers*
15 East 40th St., Dept. 186, New York 16, N. Y.

Please send me———copies of the ENCYCLOPEDIA OF CHILD GUIDANCE, at $7.50 per copy. Enclosed are $———

NAME———

ADDRESS———

PERSONNEL PSYCHOLOGY



REALISTIC RESEARCH READABLY REPORTED

PERSONNEL PSYCHOLOGY is a new "journal of applied research" covering studies in all important areas of personnel management. PERSONNEL PSYCHOLOGY is founded on the policy that the significant findings in the field can be reported to management without recourse to excessive hyperpolysyllabic sesquipedalianisms.* Readability will be achieved without the loss of either technical information or soundness by the simple process of reporting the results and implications of personnel research studies seperately from the technical procedures that produced them. The "what" in each article will be separated from the "how."

*Too many big words like "hyperpolysyllabic sesquipedalianisms."

PERSONNEL PSYCHOLOGY

is being planned to be of equal value to all workers in the field.

PERSONNEL EXECUTIVES PERSONNEL ADMINISTRATORS
PERSONNEL TECHNICIANS

SUBMIT MANUSCRIPTS TO

G. Frederic Kuder, Dept. of Psychology, Duke Univ., Durham, N. C.

Issued Quarterly Starting Spring 1948

Subscription Rate $6.00 per year

SEND SUBSCRIPTIONS TO

PERSONNEL PSYCHOLOGY INC.

1727 HARVARD ST., N.W. WASHINGTON 9, D. C.